Nichtlineare, mehrskalige künstliche Diffusion und $L^\infty(L^\infty)$-Beschränktheit bei DG-Lösungen höherer Ordnung von Erhaltungsgleichungen

D i s s e r t a t i o n

zur Erlangung des Grades eines Doktors
der Naturwissenschaften

vorgelegt von
Christian Henke
aus Kirchwalsede

genehmigt von der Fakultät für
Mathematik/Informatik und Maschinenbau
der Technischen Universität Clausthal

Tag der mündlichen Prüfung
30.10.2009

Bibliographische Information Der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Vorsitzender der Promotionskommission:	Prof. Dr. Michael Demuth, Technische Universität Clausthal
Hauptberichterstatter:	Prof. Dr. Lutz Angermann, Technische Universität Clausthal
Berichterstatter:	Prof. Dr. Gert Lube, Georg-August-Universität Göttingen

D 104

1. Auflage, 2010

ISBN: 978-3-86948-067-1

Zusammenfassung

Viele mathematische Modelle aus den Ingenieurwissenschaften, der Physik oder der Wirtschaft enthalten Konvektion und Diffusion. Häufig ist dabei die Diffusion sehr schwach ausgeprägt und kann in einigen Fällen sogar ganz vernachlässigt werden. Für die Strömung von Fluiden könnte dies bedeuten, dass eine reibungsfreie Strömung als Grenzfall von reibungsbehafteten Strömungen gedeutet wird. Diese Idee wird auch in aktuellen numerischen Lösungsverfahren für die angesprochenen Modellprobleme aufgegriffen. Die Methode wird als *Shock-capturing* bezeichnet und wird realisiert, indem ein isotroper, künstlicher Diffusionsterm eingeführt wird, der in Abhängigkeit der Gitterweite h skaliert wird. Die Notwendigkeit eines zusätzlichen Terms trägt dem Umstand Rechnung, dass die Gitterweite eines numerischen Verfahrens stets beschränkt ist und somit die Effekte von feineren Prozessen nicht berücksichtigt werden können. Der Nachteil dieses Vorgehens ist jedoch die geringere Konvergenzordnung im Vergleich zum nicht modifizierten Verfahren. Abhilfe schafft hier die Berücksichtigung eines residual basierten Diffusionsterms, der allerdings eine nichtlineare Methode zur Folge hat.

Zusammen mit einer weiteren residualbasierten, anisotropen Stabilisierung konnte in [JJS95] für eine diskrete Lösung einer Erhaltungsgleichung eine gleichmäßige Beschränktheit in der $L^\infty(L^\infty)$-Norm gezeigt werden.

In dieser Arbeit wurde ein Shock-capturing Verfahren konstruiert, das ohne anisotrope Stabilisierung ebenfalls eine gleichmäßige Beschränktheit in der $L^\infty(L^\infty)$-Norm garantiert. Die dafür vorgenommene Modifikation zerstört jedoch den residualen Charakter der Methode. Um die Aussicht auf eine höhere Konvergenzordnung zu erhalten, wurde die Idee verfolgt, die künstliche Diffusion mit einem speziell konstruierten Fluktuationsoperator P_h' nur auf feine Skalen zu projizieren, sodass die gleichmäßige Beschränktheit der Lösung erhalten bleibt.

Die in dieser Arbeit entwickelten Ergebnisse sind außerdem in der Lage, die $L^\infty(L^\infty)$-Abschätzungen der Methode aus [JJS95] bei beliebigen Polynomgraden und quasiuniformen Partitionierungen zu gewährleisten. Bisher gelang dies nur für lineare Ansatzfunktionen auf Dreiecken mit einem rechten Winkel.

Inhaltsverzeichnis

Rein gar nichts ereignet sich in der Welt, worin nicht ein Gesetz des Maximums oder Minimums zutage tritt.

(Leonard Euler)

1 Einleitung

Eine wichtige Anwendung in den Ingenieurwissenschaften und in der Physik ist die Berechnung der Strömungen von Flüssigkeiten und Gasen. Es ist heute allgemein anerkannt, dass die mathematische Beschreibung der Strömungsvorgänge durch die *Navier-Stokes Gleichungen* zusammen mit den Randbedingungen des zu umströmenden Gebiets, sämtliche realen Strömungen vollständig erfassen. Die Navier-Stokes Gleichungen sind dabei nichts anderes als Erhaltungsgleichungen für den Impuls, die Masse und die Energie im Kontinuum. In einem inkompressiblen Medium kann die Energieerhaltungsgleichung unabhängig von den beiden anderen Gleichungen gelöst werden und wird aus diesem Grund in der Literatur häufig nicht zu den Navier-Stokes Gleichungen gezählt. In diesem Fall beschreiben die Gleichungen die Entwicklung des Geschwindigkeitsfeldes und des Druckfeldes.

Lässt sich die innere Reibung des Fluids vernachlässigen, so entstehen aus den Navier-Stokes Gleichungen die *Euler Gleichungen*. Der Vorteil der allgemeinen Anwendbarkeit der Navier-Stokes Gleichungen wird jedoch durch ihre Komplexität wieder zunichte gemacht: Nur in einigen, wenigen Ausnahmefällen lassen sich diese Gleichungen analytisch lösen. Nicht einmal die Eindeutigkeit von sogenannten schwachen Lösungen der inkompressiblen, dreidimensionalen Navier-Stokes Gleichungen kann garantiert werden.

Eine weitere Fragestellung, die sich mit der Realitätsnähe der Navier-Stokes Gleichungen beschäftigt, gilt als eines der härtesten, ungelösten mathematischen Probleme unserer Zeit: Lassen die Navier-Stokes Gleichung auch Lösungen zu, die an einem Startpunkt glatt sind und bei der sich zu einem späteren Zeitpunkt die Turbulenzen so aufschaukeln, dass die Strömungsenergie unendlich wird [Fef00]? Sollte dies der Fall sein, so wären die Navier-Stokes Gleichungen ein schlechtes Modell für die Realität. Darüberhinaus ist die Konvergenz von numerischen Methoden für die Navier-Stokes Gleichungen nur bei entsprechender Glattheit der exakten Lösungen garantiert.

Da die exakte Lösung der Navier-Stokes Gleichung i.Allg. außer Reichweite liegt, ist eine numerische Approximation der Lösung häufig der einzige Ausweg um praktisch relevante Strömungsverhältnisse zu bestimmen. Allerdings offenbart das „naive“ Diskretisieren mit einem beliebigen numerischen Verfahren auch hier Schwierigkeiten. Diese direkte numerische Simulation (DNS) erfordert eine Feinheit der Diskretisierung, die in der Lage ist, auch die kleinsten auftretenden Wirbel darzustellen. Abschätzungen, die auf dem Kolmogorovschen Energiespektrum beruhen, geben für die dafür notwendige Anzahl an räumlichen und zeitlichen Freiheitsgraden eine Größenordnung von Re^3 an [Fri96, S. 107]. Re ist hierbei die dimensionslose Reynoldszahl, definiert als $\mathrm{Re} = UL/\nu$, wobei U eine charakterische Strömungsgeschwindigkeit, L eine charakteristische Längeneinheit und ν die charakteristische kinematische Viskosität darstellt. Da bei turbulenten Strömungen eine Reynoldszahl von 10^7 keine Seltenheit ist, ist die DNS auch in Zukunft kein praktikables Verfahren. Ein Ausweg ist es, die nicht aufgelösten, feinsten Skalen in einem zusätzlichen Term zu modellieren.

Das oben angesprochene „naive“ Diskretisieren der kompressiblen Navier-Stokes Gleichungen in der dimensionslosen Form

$$\begin{aligned} \partial_t u + (u \cdot \nabla)\, u - \mathrm{Re}^{-1}\, \Delta u + \nabla p &= f, \\ \nabla \cdot u &= 0, \end{aligned} \tag{1.1}$$

für die Strömungsgeschwindigkeit $u = (u_1, u_2)$ bzw. $u = (u_1, u_2, u_3)$, dem Druck p und einer von außen einwirkenden Kraft f (vgl. z.B. [MT98, S. 512 ff.] oder [EGK08, S. 222]) inklusive der Anfangs- und Randbedingungen, erweist sich auch bei moderaten Reynoldszahlen als tückisch. Als Grund dafür kristallisiert sich bei der Analysis des Verfahrens die Nichtlinearität des Terms $(u \cdot \nabla)\, u$ und die Bedingung $\nabla \cdot u = 0$ heraus.

Die spezielle Form von (1.1) erlaubt u.a. eine gewisse Modifikation der Größen U, L und ν. Solange die Reynoldszahl dabei unverändert bleibt, ergeben sich aus der gleichbleibenden Differentialgleichung (1.1), unter Beachtung der entsprechenden Skalierung, stets identische Lösungen.

Die bereits erwähnten Euler Gleichungen ergeben sich für $\mathrm{Re} \to \infty$ bzw. $\nu \to 0$:

$$\begin{aligned} \partial_t u + (u \cdot \nabla)\, u + \nabla p &= f, \\ \nabla \cdot u &= 0. \end{aligned} \tag{1.2}$$

Beim Studium von numerischen Verfahren ist es aus den genannten Gründen hilfreich, die Schwierigkeiten getrennt voneinander anzugehen und z.B. folgendes skalarwertige Modellproblem zu betrachten:

$$\partial_t u + \nabla \cdot f_s(u) - \epsilon \Delta u = f, \tag{1.3}$$

wobei die Funktion f_s eine geeignete, nichtlineare Funktion ist. Da der Begriff der Reynoldszahl in diesem vereinfachten Modell unangemessen erscheint, übernimmt der Parameter $\epsilon \approx \mathrm{Re}^{-1}$ diese Aufgabe. Der reibungsfreie Fall $\mathrm{Re} \to \infty$ wird daher durch $\epsilon \to 0$ realisiert:

$$\partial_t u + \nabla \cdot f_s(u) = f. \tag{1.4}$$

Betrachtet wird in dieser Arbeit jedoch nur die zu (1.4) gehörende *Erhaltungsgleichung*, d.h. es treten keine externen Kräfte auf $(f = 0)$.

1.1 Mathematische und physikalische Motivation

Die Näherung der reibungsbehafteten Gleichungen ((1.1) bzw. (1.3)) durch eine Gleichung ohne Reibung ((1.2) bzw. (1.4)) stellt aus mathematischer Sicht auch bei hoher Reynoldszahl bzw. kleinem ϵ einen schwerwiegenden Eingriff dar, denn es entspricht dem Streichen des Differentialoperators mit der höchsten Ordnung. Da in diesem Fall ein anderer Typ von Differentialgleichung entsteht, ändert sich i.Allg. auch das Verhalten der Lösung entscheidend. Dieses Verhalten führt zu Problemen bei der Betrachtung von schwachen Lösungen von (1.4). Eine eindeutige Lösung kann nun nicht mehr gewährleistet werden. Damit nicht genug: Man kann zeigen, dass unter diesen Lösungen solche existieren, die den zweiten Hauptsatz der Thermodynamik verletzen [AS97, S. 808 ff.]. Lassen wir hingegen für (1.4) nur solche Lösungen zu, die auch (1.3) für $\epsilon \to 0$ erfüllen (*Viskositätslösung*), so ensteht eine eindeutige Lösung, die physikalisch gutartig ist.

1.2 Numerische Verfahren

Die Idee der Viskositätslösung wird in numerischen Verfahren mit der Eigenschaft

$$h \to 0 \Rightarrow \epsilon \to 0$$

aufgegriffen. Die Addition eines künstlichen Diffusionskoeffizienten der Form

$$\epsilon = Ch|\nabla U| \tag{1.5}$$

oder

$$\epsilon = Ch^2|\nabla U|,\ C > 0, \tag{1.6}$$

mit der Gitterweite h und der Lösung des numerischen Verfahrens U, geht zurück auf [NR50] und [Sma63]. Ziel dieser Modifikation war ursprünglich nicht die Konvergenz zur Viskositätslösung, sondern die Verbreiterung des Gebiets, in dem die Lösung starken Veränderungen unterworfen ist (*Grenzschicht*), so dass diese mit der Gitterweite h aufgelöst werden kann.

Die beiden Diffusionskoeffizienten resultieren in einer nichtlinearen Diskretisierungsmethode. Lineare Verfahren ergeben sich entsprechend aus $\epsilon = Ch$ bzw. $\epsilon = Ch^2$. Eine weitere Interpretation der künstlichen Diffusion, ist die bereits angesprochene Modellierung der nicht aufgelösten, feinsten Wirbel. So ist (1.6) bekannt als *Smagorinsky-Modell* innerhalb der *Large Eddy Simulation* (LES).

Durch das Ausweiten der Grenzschicht werden zwar nichtoszillierende Lösungen erzeugt, aber der isotrope Charakter der Stabilisierung erzeugt auch Diffusion senkrecht zur Stromlinienrichtung, so dass auch in dieser Richtung die Grenzschichten ausgeweitet werden. Als Folge entsteht ein Fehler, der im günstigsten Fall von der Ordnung $\mathcal{O}(h)$ ist.

Im Gegensatz dazu fügt die *Stromliniendiffusionsmethode* zusätzliche Stabilisation in Stromlinienrichtung hinzu [HB79], [HT84]. Dies geschieht mittels eines gewichteten Residualansatzes innerhalb des Stabilisierungsterms, der eine Fehlerordnung von $\mathcal{O}(h^{k+1/2})$ garantiert [Joh87, Theorem 9.2]. Weitere Beiträge finden sich in [HFM86], [HM86a] und [HFM87]. Eine Modifikation unter Hinzunahme eines zweiten, nichtlinearen Stromliniendiffusionsterms, mit Stabilisierung in Richtung ∇U (*Shock-capturing*), ist in [HMM86], [HM86b], [JS86a] und [JS87] beschrieben. Die resultierende nichtlineare Methode enthält durch die Projektion eine Normalisierung mit $|\nabla U|^{-2}$. Der Austausch des Normalisierungsterms mit einer größeren Potenz von h erlaubt in [Sze89a] den Nachweis der Konvergenzordnung von $\mathcal{O}(h^{k+1/2})$. Letztgenannte Modifikation erlaubt die Interpretation des Shock-capturing Terms als residualbasierte, isotrope künstliche Diffusion. Siehe hierzu auch [JSH90], [Sze91] und [JJS95].

Diese Arbeit geht der Frage nach, ob die Discontinuous-Galerkin-Methode aus [JJS95] derart modifiziert werden kann, dass bei Vernachlässigung des Stromliniendiffusionsterms die $L^\infty(L^\infty)$-Beschränkung der diskreten Lösung erhalten bleibt. Diese Fragestellung schöpft Motivation aus der Tatsache, dass die Discontinuous-Galerkin-Methode von Haus aus mit einem Fehler der Ordnung $\mathcal{O}(h^{k+1/2})$ - bzgl. einer Norm, die auch die Ableitung in Stromlinienrichtung umfasst - aufwarten kann [JP86], so dass die Stromliniendiffusionsmethode vor diesem Hintergrund keine zusätzliche Stabilität bringt. Es zeigt sich, dass die Stromliniendiffusionsmethode in [JJS95] für die $L^\infty(L^\infty)$-Abschätzung vernachlässigbar ist. Die hier vorgestellte Modifikation benötigt zwar keine Diffusion in Stromlinienrichtung mehr, zerstört allerdings

die residuale Struktur der Methode, so dass die Konvergenzordnung nur noch identisch ist mit der klassischen künstlichen Diffusion. Um trotzdem Aussicht auf eine zufriedenstellende Konvergenzordnung zu haben, wirkt der Diffusionsterm mittels Projektion nur auf die dazugehörigen feinen Skalen. In [Hei07] konnte mit dieser Strategie, allerdings bei linearer Diffusion und ohne Berücksichtigung der $L^\infty(L^\infty)$-Frage, eine höhere Konvergenzordnung erzielt werden.

1.3 Gliederung der Arbeit

Diese Arbeit umfasst acht Kapitel. Das direkt an die Einleitung anschließende Kapitel gibt einen Überblick über grundlegende Definitionen und Aussagen, die zum Verständnis dieser Arbeit benötigt werden. Insbesondere wird ein Beweis des Spursatzes für *Sobolev-Räume* vorgelegt, der für die Analysis des Discontinuous-Galerkin-Verfahrens essentiell ist.

Im 3. Kapitel werden zwei Lösungstheorien präsentiert, die mittels der Viskositätsmethode eine eindeutige schwache Lösung für (1.3) gewährleisten. Die bereits angesprochene Problematik der Typänderung der Differentialgleichung wird durch die Einführung einer verallgemeinerten Randbedingung berücksichtigt.

Das 4. Kapitel erläutert die Finite-Elemente-Methode und legt mit einigen Projektions- und Interpolationsfehlern sowie mit inversen Ungleichungen die Grundlagen für die $L^\infty(L^2)$- bzw. $L^\infty(L^\infty)$-Abschätzungen der diskreten Lösung des Discontinuous-Galerkin-Verfahrens aus Kapitel 5. Die dabei auftretenden technischen Schwierigkeiten werden überwunden, indem Ideen aus der Theorie der *hierarchischen modalen Basisfunktionen* mit Hilfe von *mass lumping* auf nodale Basisfunktionen übertragen werden. Es zeigt sich, dass auf diesem Weg ein diskreter *Fluktuationsoperator* definiert werden kann, der innerhalb eines Shock-capturing Terms die $L^\infty(L^\infty)$-Abschätzung gewährleisten kann. Eine anisotrope Stabilisierung ist in diesem Fall nicht mehr notwendig.

Die in diesem Kapitel bewiesenen Theoreme sind außerdem in der Lage, die $L^\infty(L^\infty)$-Abschätzung der Methode [JJS95, (2.7)] bei beliebigen Polynomgraden und quasiuniformen Partitionierungen zu gewährleisten. Bisher gelang dies nur für lineare Ansatzfunktionen auf Dreiecken mit einem rechten Winkel [Sze89a].

In Kapitel 6 werden Konvergenzordnungsaussagen der nichtlinearen Discontinuous-Galerkin-Methode mit Fluktuationsoperator und der Methode aus [JJS95, (2.7)] vorgestellt. Diese Untersuchungen erfolgen für das lineare, homogene Problem (1.4) und berücksichtigen sowohl die Abhängigkeit der Gitterweite h als auch die Abhängigkeit des Polynomgrades.

Die numerische Umsetzung des Verfahrens wurde mit der Finite-Elemente Bibliothek *deal.II* [BHK] vorgenommen und getestet. Entsprechende Beispiele sind im 7. Kapitel zu finden.

Das Kapitel 8 schließt die Arbeit mit einer zusammenfassenden Diskussion und einem Ausblick. Hier werden die Aussagen zur Konvergenzordnung im Hinblick auf die erzielten numerischen Ergebnisse kommentiert und mögliche Erweiterungen sowie offene Fragestellungen diskutiert.

1.4 Danksagung

Die in dieser Arbeit präsentierten Ergebnisse entstammen im Wesentlichen meiner fünfjährigen Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Mathematik an der TU Clausthal.

An dieser Stelle möchte ich mich ganz herzlich bei meinem Doktorvater Prof. Dr. Lutz Angermann für die gute Betreuung, die interessanten Diskussionen und alternativen Sichtweisen bedanken.

Prof. Dr. Gert Lube danke ich für die freundliche Übernahme des Korreferates und Prof. Dr. Ulrich Mertins für die überaus motivierende Begleitung während meiner ersten Studiensemester.

Zu guter Letzt möchte ich allen Kirchwalsedern danken, die speziell in diesem Jahr reges Interesse am Verlauf der Promotion ihres aktuellen Schützenkönigs gezeigt haben.

2 Funktionalanalytische Grundlagen

Für die Betrachtung der hyperbolischen Gleichungen ist es notwendig, sofern die für diese Gleichungen typischen unstetigen Lösungen berücksichtigt werden sollen, auf eine schwache Formulierung der Differentialgleichung zurückzugreifen. Die dafür nötigen Hilfsmittel sollen in der üblichen Notation kurz vorgestellt werden.

2.1 Funktionenräume

Es bezeichne $C^l(G,\mathbb{K})$, $l \in \mathbb{N}_0$ die l-mal stetig differenzierbaren Abbildungen von einem *Gebiet* $G \subset \mathbb{K}^m$, $m \in \mathbb{N}$ in den Körper $\mathbb{K}$. $C_0^l(G,\mathbb{K})$ enthält Funktionen aus $C^l(G,\mathbb{K})$, deren Support kompakt ist. Für $0 < \theta \leq 1$ ist $C^{l,\theta}(\overline{G},\mathbb{K})$ der Unterraum von $C^l(\overline{G},\mathbb{K})$, bestehend aus Funktionen, so dass $\partial^\alpha u$, $0 \leq \alpha \leq l$ für jede Komponente u mit dem Exponenten θ *Hölderstetig* ist:

$$|\partial^\alpha u(x) - \partial^\alpha u(y) \leq C|x-y|^\theta, \quad x,y \in G.$$

Der *Lebesgue-Raum* $L^p(G,\mathbb{K})$ mit $p \in [1,\infty)$ enthält die messbaren Funktionen $u : G \to \mathbb{K}$ für die im Lebesgueschen Sinne gilt: $\|u\|^p_{L^p(G,\mathbb{K})} = \int_G |u|^p dx < \infty$. Für $p = \infty$ ist $L^\infty(G,\mathbb{K})$ der Raum der wesentlich beschränkten Funktionen, d.h. $u : G \to \mathbb{K}$ ist messbar und fast überall beschränkt. In diesem Fall wird die Norm als die kleinste dieser Schranken definiert:

$$\|u\|_{L^\infty(G,\mathbb{K})} = \operatorname*{ess\,sup}_{x\in G}|u(x)| = \inf\{K > 0 : |u(x)| \leq K \text{ fast überall in } G\}.$$

Mit der eingeführten Notation bedeutet $\|u-v\|_{L^p(G,\mathbb{K})} = 0$, dass $u(x) = v(x)$ nur fast überall (f.ü.) auf G gilt. Folglich werden die Elemente von $L^p(G,\mathbb{K})$ als Äquivalenzklassen von Funktionen aufgefasst. Wenn trotzdem von Funktionen aus $L^p(G,\mathbb{K})$ die Rede ist, so ist in diesem Fall stets ein Repräsentant der Äquivalenzklasse gemeint. Aus diesem Grund werden solche Funktionen, die sich nur auf einer Menge vom Maß Null unterscheiden, üblicherweise miteinander identifiziert.

Der Vektor $\alpha = (\alpha_1,\ldots,\alpha_m)^T \in \mathbb{N}_0^m$ heißt *Multiindex*. Mit $x = (x_1,\ldots,x_m)^T \in \mathbb{R}^m$ sind die Schreibweisen $\partial^\alpha = \partial_1^{\alpha_1}\cdots\partial_m^{\alpha_m}$, $\alpha! = \alpha_1!\cdots\alpha_m!$, $x^\alpha = x_1^{\alpha_1}\cdots x_m^{\alpha_m}$ und $|\alpha| = \sum_{i=1}^m \alpha_i$ üblich. Für ein $l \in \mathbb{N}_0$, $p \in [1,\infty)$ und einen Multiindex α definiert bei schwacher Differentiation und geeignetem u

$$\|u\|_{l,p,G} = \left(\sum_{|\alpha|\leq l} \|\partial^\alpha u\|^p_{L^p(G,\mathbb{K})}\right)^{1/p}$$

eine Norm. Im Fall $p = \infty$ ist

$$\|u\|_{l,\infty,G} = \max_{|\alpha|\leq l}\|\partial^\alpha u\|_{L^\infty(G,\mathbb{K})}$$

die entsprechende Norm. Die dazugehörigen Sobolev-Räume $W^{l,p}(G,\mathbb{K})$ enthalten die Funktionen, deren entsprechende Norm endlich ist. Im Fall $p=2$ existieren jeweils die kanonischen Innenprodukte

$$(u,v)_{l,G}=\sum_{|\alpha|\le l}\int_G \partial^\alpha u\,\partial^\alpha v\,dx.$$

Sobolev-Räume mit einem reellen Index, etwa $l+\theta,\ 0<\theta<1$ für $1\le p<\infty$ werden mit Hilfe der K-Methode der Interpolationstheorie definiert. Entsprechende Definitionen und Eigenschaften finden sich in [BS08, Chapter 14] und [AF03, Chapter 7]. Mit diesen Voraussetzungen soll gelten

$$W^{l+\theta,p}(G)=[W^{l,p}(G),W^{l+1,p}(G)]_{\theta,p}, \tag{2.1}$$

$$\|u\|_{l+\theta,p,G}=\left(\int_0^\infty \left(t^{-\theta}K(t,u)\right)^p\frac{dt}{t}\right)^{1/p} \tag{2.2}$$

und

$$\|u\|_{l+\theta,\infty,G}=\sup_{0<t<\infty} t^{-\theta}K(t,u). \tag{2.3}$$

Im Zuge dieser Vereinbarungen und der Normäquivalenz aus [BS08, 14.0.1, Theorem 14.2.3] erscheint es sinnvoll, die dazugehörigen Halbnormen für $p<\infty$ durch die Ausdrücke

$$|u|_{l,p,G}=\begin{cases}\left(\sum_{|\alpha|=l}\|\partial^\alpha u\|^p_{L^p(G,\mathbb{K})}\right)^{1/p}, & l\in\mathbb{N},\\ \left(\sum_{|\alpha|=\lfloor l\rfloor}\int_G\int_G\frac{|\partial^\alpha u(x)-\partial^\alpha u(y)|^p}{|x-y|^{d+\theta p}}dx\,dy\right)^{1/p}, & l=\lfloor l\rfloor+\theta,\ \theta>0\end{cases} \tag{2.4}$$

zu definieren. Für $p=\infty$ definiere analog (vgl. [AF75, S. 7.49])

$$|u|_{l,\infty,G}=\begin{cases}\max\limits_{|\alpha|=l}\|\partial^\alpha u\|_{L^\infty(G,\mathbb{K})}, & l\in\mathbb{N},\\ \max\limits_{|\alpha|=\lfloor l\rfloor}\operatorname{ess\,sup}\limits_{\substack{x,y\in G\\ x\neq y}}\frac{|\partial^\alpha u(x)-\partial^\alpha u(y)|}{|x-y|^\theta}, & l=\lfloor l\rfloor+\theta,\ \theta>0.\end{cases} \tag{2.5}$$

Besteht $\mathbb{K}$ aus dem Körper der reellen Zahlen, so sollen die Konventionen $C^l(G)=C^l(G,\mathbb{R})$, $C^{l,\theta}(\overline{G})=C^{l,\theta}(\overline{G},\mathbb{R})$, $L^p(G)=L^p(G,\mathbb{R})$ und $W^{l,p}(G)=W^{l,p}(G,\mathbb{R})$ gelten. Ausführliche Darstellungen und alternative Definitionen der Sobolev-Räume sind nachzulesen in [AF03] und [Wlo82]. Einführende Darstellungen finden sich in [Alt06] und [War99].

Eine Funktion $f\in L^1(G)$ ist von *beschränkter Variation* in G, wenn

$$\int_G|\mathcal{G}f|dx=\sup_{\substack{|g(x)|\le 1\\ x\in G}}\left\{\int_G f\ \operatorname{div}\ g\,dx : g=(g_1,\ldots,g_n)^T\in C_0^1(G)^n\right\}<\infty$$

gilt. Der Raum $BV(G)$ wird definiert als die Menge der Funktionen aus $L^1(G)$ mit beschränkter Variation (vgl. [Giu84]).

Die *Kardinalzahl* einer Menge $\mathcal{M}$ wird mit $\#\mathcal{M}$ bezeichnet. Ist $v=(v_i)_{i\in\mathcal{I}}\in\mathbb{R}^{\#\mathcal{I}}$ und $\mathcal{I}$ eine Indexmenge, so werden wie üblich folgende Normen definiert:

$$\|v\|_{l^p(\mathcal{I})}=\left(\sum_{i\in\mathcal{I}}|v_i|^p\right)^{1/p}\ \text{für } p\in(0,\infty)\ \text{und}\ \|v\|_{l^\infty(\mathcal{I})}=\max_{i\in\mathcal{I}}\{|v_i|\}.$$

Im Fall $p = 2$ existiert das l^2-Innenprodukt

$$(v, w)_{l^2(\mathcal{I})} = \sum_{i \in \mathcal{I}} v_i w_i,\ w = (w_i)_{i \in \mathcal{I}}.$$

Erfolgt die Summierung mit der Indexmenge $\mathcal{I} = \{0, 1, 2, \ldots, n-1\}$ oder $\mathcal{I} = \{1, 2, \ldots, n\}$ mit $n = \#\mathcal{I}$, so findet die abkürzende Schreibweise $\|\cdot\|_{l^p}$ bzw. $(\cdot, \cdot)_{l^p}$ Anwendung. Für zwei unterschiedliche Indizes $p_1 < p_2$ gelten nach [Gas70, S. 28] die Abschätzungen

$$\|v\|_{l^{p_2}} \leq \|v\|_{l^{p_1}} \leq n^{(1/p_1 - 1/p_2)} \|v\|_{l^{p_2}}. \tag{2.6}$$

Bei Funktionen auf einem Gebiet G definiert

$$\|f\|_{l^p(\mathcal{M})} = \left\{ \sum_{x \in \mathcal{M}} |f(x)|^p \right\}^{1/p}$$

eine diskrete Norm auf einer Punktmenge $\mathcal{M}$. $\|f\|_{l^\infty(\mathcal{M})}$ wird analog definiert.

Für $A \in \mathbb{R}^{n,n}$, $n \in \mathbb{N}$ bezeichnet $\|A\|_{l^p}$, $1 \leq p \leq \infty$ die von $\|x\|_{l^p}$, $x \in \mathbb{R}^n$ induzierte Norm und $\|A\|_{\max} = \max\limits_{1 \leq i,j \leq n} |a_{ij}|$.

2.2 Geometrische Voraussetzungen und der Spuroperator

Da die Eigenschaften von Funktionen aus den Sobolev-Räumen $W^{l,p}(G)$ stark von den Eigenschaften des Randes ∂G abhängen, ist bei der Betrachtung des zugrunde liegenden Gebietes besondere Aufmerksamkeit bei der notwendigen Glattheit des Randes gefordert (vgl. z.B. [Gri85], [Alt06, A6] und [AF03]).

Definition 2.2.1 Sei $G \subset \mathbb{R}^d$, $d \in \mathbb{N}$ ein *beschränktes Gebiet*. G besitzt einen C^m*-Rand* [$C^{m,1}$*-Rand*], falls sich ∂G durch endlich viele offene Mengen $U^1, \ldots, U^r$ überdecken lässt, so dass $\partial G \cap U^j$ für $j = 1, \ldots, r$ der Graph einer Funktion aus C^m [$C^{m,1}$] ist und $G \cap U^j$ auf jeweils einer Seite dieses Graphen liegt. D.h. es gibt für $1 \leq j \leq r$ ein euklidisches Koordinatensystem $e_1^j, \ldots, e_d^j$ in $\mathbb{R}^d$, eine Abbildung $\varphi^j : \mathbb{R}^{d-1} \to \mathbb{R}$, $\varphi^j \in C^m(F^j(\overline{B}_{d-1}))$ [$\varphi^j \in C^{m,1}(F^j(\overline{B}_{d-1}))$], so dass für die bijektive Funktion $\psi^j : \overline{B} \to U^j$ mit

$$\psi^j(x) = \sum_{i=1}^{d-1} F^j(x)_i e_i^j + (F^j(x)_d + \varphi^j(F^j(x)_1, \ldots, F^j(x)_{d-1})) e_d^j,$$

$F^j(x) = diag(r^j, \ldots, r^j, h^j)x$, $r^j, h^j > 0$ die Identitäten

$$\begin{aligned} U^j \cap \partial G &= \psi^j(B_0), \\ U^j \cap G &= \psi^j(\{y \in B : y_n > 0\}), \end{aligned}$$

gelten sollen. Dabei ist $B = \{y \in \mathbb{R}^d : \|y\|_{l^\infty} < 1\}$, $B_{d-1} = \{y \in \mathbb{R}^{d-1} : \|y\|_{l^\infty} < 1\}$ und $B_0 = B_{d-1} \times \{0\}$. Handelt es sich um einen $C^{0,1}$-Rand, so wird auch von einem *Lipschitz-Rand* gesprochen.

Die Tangentialvektoren ergeben sich zu

$$\tau_s(y) = \frac{\partial}{\partial y_s}\left(\sum_{i=1}^{d-1} y_i e_i^j + \varphi^j(y')e_d^j\right) = e_s^j + \frac{\partial}{\partial y_s}\varphi^j(y')e_d^j,\ y \in F^j(B_0), \tag{2.7}$$

mit $y' = (y_1, \ldots, y_{d-1})^T$ für $1 \leq s \leq d-1$. Der zu allen Tangentialvektoren senkrecht stehende äußere Normalenvektor ist

$$n(y) = \left(1 + \|\nabla\varphi^j\|_{l^2}^2\right)^{-1/2}\left(\sum_{i=1}^{d-1}\frac{\partial}{\partial y_i}\varphi^j(y')e_i^j - e_i^d\right),\ y \in F^j(B_0). \tag{2.8}$$

Alle Ableitungen sind im schwachen Sinn zu verstehen und existieren für $\varphi^j \in C^1(F^j(\overline{B}_{d-1}))$ und $\varphi^j \in C^{0,1}(F^j(\overline{B}_{d-1}))$, denn es gilt nach [Alt06, Satz 8.5 (2)] für ein *Lipschitz-Gebiet* G die Einbettung $C^{0,1}(\overline{G}) \subset W^{1,\infty}(G)$ und somit $\nabla\varphi^j \in L^\infty(F^j(B_0))$.

Zur Interpretation der Integration auf ∂G bzw. zur Definition der Sobolev-Räume über den Rand, ist es notwendig die lokalen Eigenschaften der Definition 2.2.1 auf den gesamten Rand zu übertragen. Dies geschieht mittels einer *Partitionierung der Eins* bezüglich der offenen und endlichen Überdeckung U^j (vgl. [Alt06, 2.19, A 6.3]). Durch Hinzufügen einer weiteren offenen Menge U^0 ist es möglich, ganz G zu überdecken. Unter diesen Voraussetzungen existieren

$$\omega^j : \mathbb{R}^d \to \mathbb{R},\ \omega^j \in C_0^\infty(U^j),\ \omega^j \geq 0$$

und $\sum_{i=0}^r \omega^j(x) = 1$ für $x \in G$.

Ferner ist ψ^j in der Form $\psi^j : \mathbb{R}^{d-1} \times \{0\} \supset B_{d-1} \times \{0\} \to U^j \cap \partial G \subset \mathbb{R}^d$ eine *Parameterdarstellung* für das Oberflächenstück $U^j \cap \partial G$, dessen Oberfläche sich mit Hilfe von

$$G = J_{\psi^j}^T J_{\psi^j}, \quad J_{\psi^j} = \left(\frac{\partial}{\partial x_n'}\psi_m^j\right)_{\substack{1\leq m\leq d\\ 1\leq n\leq d-1}}$$

bestimmen lässt durch

$$s_{\psi^j} = \int_{B_{d-1}} (\det\ G(x'))^{1/2}\ dx' = \int_{B^{d-1}} g(x')\,dx'.$$

Durch Nachrechnen lässt sich $g(x') = \left(1 + \|\nabla\varphi^j(x')\|_{l^2}^2\right)^{1/2}$ bestätigen. Das Randintegral für f kann nun mit Hilfe von $\operatorname{supp} f\omega^j \subset U^j$ ausgedrückt werden als

$$\begin{aligned}\int_{\partial G} f(s)\,ds &= \sum_{j=1}^r \int_{\partial G} f(s)\omega^j(s)\,ds = \sum_{j=1}^r \int_{U^j\cap\partial G} f(s)\omega^j(s)\,ds\\ &= \sum_{j=1}^r \int_{\psi^j(B_0)} \left(f\omega^j\right)\circ\psi^j(x',0)\,ds = \sum_{j=1}^r \int_{B_0} \left(f\omega^j\right)\circ\psi^j(x',0)\,ds_{\psi_j}\\ &= \sum_{j=1}^r \int_{\mathbb{R}^{d-1}} \left(f\omega^j\right)\circ\psi^j(x',0)g(x')\,dx.'\end{aligned} \tag{2.9}$$

Wie bereits angedeutet gilt $\nabla\varphi^j \in C^{m-1}(F^j(\overline{B}_{d-1}))\,[\nabla\varphi^j \in W^{m,\infty}(F^j(B_{d-1}))]$ und somit

$$1 \leq g^{(l)}(x') \leq C, \text{ fast überall in } \mathbb{R}^{d-1},\ 0 \leq l \leq m-1\,[0 \leq l \leq m]. \tag{2.10}$$

Mit diesen Hilfsmitteln kann die L^p-Norm für $1 \leq p < \infty$ definiert werden:

$$\begin{aligned}\int_{\partial G} |u|^p \, ds &= \sum_{j=1}^{r} \int_{\partial G} |u|^p \omega^j \, ds \\ &\underset{(2.9)}{=} \sum_{j=1}^{r} \int_{\mathbb{R}^{d-1}} |u \circ \psi^j(x',0)|^p \omega^j \circ \psi^j(x',0)\, g(x') \, dx' \\ &= \sum_{j=1}^{r} \|u \circ \psi^j\|^p_{0,p,\omega^j \circ \psi^j g, \mathbb{R}^{d-1}} = \|u\|^p_{0,p,\partial G}.\end{aligned} \tag{2.11}$$

Die Definition der Sobolev-Norm folgt entsprechend:

$$\begin{aligned}\sum_{|\alpha| \leq l} \|\partial^\alpha u\|^p_{0,p,\partial G} &= \sum_{|\alpha| \leq l} \sum_{j=1}^{r} \|\partial^\alpha u \circ \psi^j\|^p_{0,p,\omega^j \circ \psi^j g, \mathbb{R}^{d-1}} \\ &= \sum_{j=1}^{r} \|u \circ \psi^j\|^p_{l,p,\omega^j \circ \psi^j g, \mathbb{R}^{d-1}} = \|u\|^p_{l,p,\partial G}.\end{aligned} \tag{2.12}$$

Im Fall $p = \infty$ lassen sich die Normdefinitionen analog zum Gebiet vornehmen:

$$\|u\|_{0,\infty,\partial G} = \operatorname*{ess\,sup}_{s \in \partial G} |u(s)|, \tag{2.13}$$

$$\|u\|_{l,\infty,\partial G} = \max_{|\alpha| = l} \operatorname*{ess\,sup}_{s \in \partial G} \|\partial^\alpha u\|_{0,\infty,\partial G}. \tag{2.14}$$

Alle hier vorgestellten Definitionen, die auf dem lokalen Koordinatensystem der Definition 2.2.1 beruhen, sind unabhängig von der lokalen Zerlegung und Darstellung des Randes [Alt06, A 6.5].

Wird im Folgenden ohne weitere Anmerkung von einem Gebiet G gesprochen, so ist stets ein Gebiet mit Lipschitz-Rand gemeint, insbesondere gilt dann die Formel der partiellen Integration und der *Satz von Gauss* [Alt06, A 6.8].

Die Frage nach der Existenz der Randintegrale ist für den Fall $u \in C(\overline{\Omega})$ leicht zu beantworten, da $u|_{\partial\Omega}$ in kanonischer Weise durch Restriktion entsteht. Allerdings ist für $W^{l,p}(\Omega) \subset C(\overline{\Omega})$ die Bedingung $l > d/p$ (vgl. [EG04, Corollary B.46]) erforderlich. Somit scheint es für $l \leq d/p$ keinen Sinn zu machen von Randwerten zu sprechen, da nach Definition der Lebesgue-Räume miteinander identifizierte Funktionen auf $\partial\Omega$ verschiedene Werte annehmen können. Der nachstehende Spursatz für Gebiete garantiert jedoch für $l > 1/p$ die Existenz der Randintegrale. In der Literatur sind verschiedene Darstellungen in unterschiedlicher Allgemeinheit zu finden, z.B. in [Gri85, Theorem 1.5.1.2, Theorem 1.5.2.1], [Wlo82, Satz 8.7]. Die in dieser Arbeit gewählte Formulierung fußt auf einem Spursatz für $\tilde{\gamma} : \mathbb{R}^d \to \mathbb{R}^{d-1}$ und verwendet im Beweis sogenannte lineare, stetige *Erweiterungsoperatoren* $E : W^{l,p}(G) \to W^{l,p}(\mathbb{R}^d)$, deren Existenz in Abhängigkeit der Randglätte gewährleistet werden kann. Die Definition von $\tilde{\gamma} : \mathbb{R}^d \to \mathbb{R}^{d-1}$ geschieht wie bereits angekündigt durch Restriktion:

$$(\tilde{\gamma} u)(x') = u(x', 0), \quad x' \in \mathbb{R}^{d-1}, \tag{2.15}$$

falls $u \in C(\mathbb{R}^d)$.

Die notwendige Variablensubstitution, um den Spursatz auf Gebiete zu übertragen ist durch die Definition 2.2.1 bereits festgelegt.

Theorem 2.2.2 *Sei $1 \leq p < \infty$ und $l \in \mathbb{R}$, $l > 1/p$. Dann existiert ein linearer, stetiger Operator $\tilde{\gamma} : W^{l,p}(\mathbb{R}^d) \to W^{l-1/p,p}(\mathbb{R}^{d-1})$.*

Beweis Unter den gegebenen Voraussetzungen ist die Existenz eines $C > 0$ mit

$$\|\tilde{\gamma}u\|_{l-1/p,p,\mathbb{R}^{d-1}} \leq C\|u\|_{l,p,\mathbb{R}^d} \tag{2.16}$$

zu zeigen. Für $p = 1$ erfüllt [AF03, 4.12 (4)] die Aufgabe. Im Fall $l \in \mathbb{N}$, $1 < p < \infty$ siehe [AF03, Theorem 7.39] und für $l \in \mathbb{R} \setminus \mathbb{N}$, $1 < p < \infty$ betrachte [AF03, Theorem 7.43] oder auch [BL76, Definition 6.2.2, Theorem 6.2.3 und Theorem 6.2.4 (10)]. □

Definition 2.2.3 (*starker m-Erweiterungsoperator*) Eine lineare Abbildung

$$E : W^{l,p}(G) \to W^{l,p}(\mathbb{R}^d)$$

wird starker m-Erweiterungsoperator für G genannt, falls für jedes p, $1 \leq p < \infty$ und jedes $l \in \mathbb{N}_0$, $0 \leq l \leq m \in \mathbb{N}_0$ eine Konstante $C = C(m)$ existiert, so dass für jedes $u \in W^{l,p}(G)$ die Bedingungen

$$Eu(x) = u(x) \text{ fast überall in } G, \tag{2.17}$$

$$\|Eu\|_{l,p,\mathbb{R}^d} \leq C\|u\|_{l,p,G}. \tag{2.18}$$

erfüllt sind.

Definition 2.2.4 (*totaler Erweiterungsoperator*) E wird totaler Erweiterungsoperator für G genannt, falls E ein starker m-Erweiterungsoperator für G und jedes m ist.

Bemerkung 2.2.5 Ein totaler Erweiterungsoperator garantiert $Eu(x) = u(x)$ für alle $x \in \overline{G}$ (vgl. [AF03, S. 5.17]).

Lemma 2.2.6 *Sei $G \subset \mathbb{R}^d$ mit einem C^m-Rand. Dann existiert ein starker m-Erweiterungsoperator E für G.*

Beweis [AF03, Theorem 5.22]. □

Lemma 2.2.7 *Sei $G \subset \mathbb{R}^d$ mit Lipschitz-Rand. Dann existiert ein totaler Erweiterungsoperator E für G.*

Beweis [AF03, Theorem 5.24] bzw. [Ste70, Chapter 6]. □

Zum Beweis des angekündigten Spursatzes wird das folgende Lemma benötigt:

Lemma 2.2.8 (Faà di Bruno's Formel) *Für einen Multiindex α und zwei Funktionen*

$$f : G \subset \mathbb{R}^m \to \mathbb{R}^n,\ g : D \subset \mathbb{R}^n \to \mathbb{R},\ m, n \in \mathbb{N}$$

deren Ableitungen bis zur Ordnung $|\alpha|$ existieren, gilt mit $P = \mathbb{N}_0^m \setminus \{(0, \ldots, 0)\}$ Faà di Bruno's Formel

$$\partial^\alpha (g \circ f)(x) = \alpha! \sum_{|\beta| \leq |\alpha|} (\partial^\beta g)(f(x)) \sum_{\substack{a : P \to \mathbb{N}_0^n \\ \sum_{\gamma \in P} a(\gamma) = \beta \\ \sum_{\gamma \in P} |a(\gamma)|\gamma = \alpha}} \prod_{\gamma \in P} \frac{1}{a(\gamma)!} \left[\frac{(\partial^\gamma f)(x)}{\gamma!} \right]^{a(\gamma)}. \tag{2.19}$$

Beweis [Gzy86]. Vgl. auch [Dix02, Lemma 2.1]. □

Theorem 2.2.9 *Es sei G ein* Gebiet *mit einem C^m-Rand [$C^{m,1}$-Rand], $m \in \mathbb{N}$. Dann existiert für $1 \leq p < \infty$ und $1/p < l$, $l \in \mathbb{R}$ ein linearer, stetiger* Spuroperator

$$\gamma : W^{l,p}(G) \to W^{l-1/p,p}(\partial G), \quad l \leq m\,[l \leq m+1]. \tag{2.20}$$

Ist G ein Gebiet mit Lipschitz-Rand, so gilt

$$\gamma u = u|_{\partial G}, \tag{2.21}$$

falls $u \in W^{l,p}(G) \cap C(\overline{G})$.

Beweis Nach Lemma 2.2.6 [2.2.7] existiert ein starker m-Erweiterungsoperator [totaler Erweiterungsoperator] E. Aus der Definition des Randintegrals folgt

$$\begin{aligned}
\|\tilde{\gamma}(Eu)\|^p_{l-1/p,p,\partial G} &= \sum_{j=1}^{r} \|\tilde{\gamma}(Eu)\circ\psi^j\|^p_{l-1/p,p,\omega^j\circ\psi^j g,\mathbb{R}^{d-1}} \\
&\underset{(2.10),\,\omega^j\leq 1}{\leq} \sum_{j=1}^{r} \|\tilde{\gamma}(Eu)\circ\psi^j\|^p_{l-1/p,p,\mathbb{R}^{d-1}} \\
&\underset{(2.16)}{\leq} \sum_{j=1}^{r} \|Eu\circ\psi^j\|^p_{l,p,\mathbb{R}^d} = \sum_{j=1}^{r} \|Eu\circ\psi^j\|^p_{l,p,B} \\
&= \sum_{j=1}^{r}\sum_{i=0}^{l} |Eu\circ\psi^j|^p_{i,p,B}.
\end{aligned}$$

Auf Grund der Voraussetzungen an den Rand von G gilt für $i = |\alpha| \leq m$: $|\partial^\alpha \psi^j_s(x)| \leq C$ $[i = |\alpha| \leq m+1 : |\partial^\alpha \psi^j_s(x)| \leq C$ f. ü. auf B, denn [Alt06, Satz 8.5 (2)] $\Rightarrow \partial^\alpha \psi^j_s \in L^\infty(B)]$ mit $1 \leq s \leq d$ und $i = 0, \ldots, l$. Aus Faà di Bruno's Formel (2.19) folgt somit

$$\hat{\partial}^\alpha \left(w\circ\psi^j\right)(\hat{x}) \leq C \sum_{\beta\leq|\alpha|} \left(\partial^\beta w\circ\psi^j(\hat{x})\right)$$

und mit $\det(\frac{\partial}{\partial\hat{x}}\psi^j) = 1$

$$\begin{aligned}
\sum_{i=0}^{l} |w\circ\psi^j|^p_{i,p,B} &= \sum_{i=0}^{l}\sum_{|\alpha|=i} \|\hat{\partial}^\alpha w\circ\psi^j\|^p_{0,p,B} \\
&= \sum_{i=0}^{l}\sum_{|\alpha|=i} \| \left(\hat{\partial}^\alpha w\circ\psi^j\right) |\det(\frac{\partial}{\partial\hat{x}}\psi^j)|^{1/p}\|^p_{0,p,B} \\
&\leq C\sum_{i=0}^{l}\sum_{|\alpha|=i}\sum_{|\beta|\leq|\alpha|} \| \left(\partial^\beta w\circ\psi^j\right) |\det(\frac{\partial}{\partial\hat{x}}\psi^j)|^{1/p}\|^p_{0,p,B} \\
&\leq C\sum_{i=0}^{l}\sum_{|\alpha|=i}\sum_{|\beta|\leq i} \|\partial^\beta w\|^p_{0,p,U^j} \\
&= C\sum_{i=0}^{l}\sum_{|\beta|\leq i} \|\partial^\beta w\|^p_{0,p,U^j} = C\sum_{i=0}^{l} \|w\|^p_{i,p,U^j} \leq C\|w\|^p_{l,p,U^j}.
\end{aligned}$$

Insgesamt ergibt sich nun die erste Behauptung

$$\begin{aligned}\|\tilde{\gamma}(Eu)\|^p_{l-1/p,p,\partial G} &\leq C\sum_{j=1}^{r}\|Eu\|^p_{l,p,U^j} \leq C\sum_{j=1}^{r}\|Eu\|^p_{l,p,\mathbb{R}^d}\\ &\leq C\|Eu\|_{l,p,\mathbb{R}^d} \leq C\|u\|^p_{l,p,G}.\end{aligned}$$

Damit wurde die Transformationskette

$$\begin{aligned}W^{l,p}(G) &\xrightarrow{E} W^{l,p}(\mathbb{R}^d) \xrightarrow{\omega^j} W^{l,p}(U^j) \overset{\left(\psi^j\right)^{-1}}{\rightarrow} W^{l,p}(B)\\ &\xrightarrow{Id} W^{l,p}(\mathbb{R}^d) \xrightarrow{\tilde{\gamma}} W^{l,p}(\mathbb{R}^{d-1}) \xrightarrow{Id} W^{l,p}(B_0)\\ &\xrightarrow{\psi^j} W^{l,p}(U^j\cap\partial G) \subset W^{l,p}(\partial G)\end{aligned}$$

beschritten. Dies liefert die Definition des Spuroperators:

$$\gamma u = \sum_{j=1}^{r}\psi^j\circ\tilde{\gamma}\circ\left(\psi^j\right)^{-1}\circ\left(\omega^j E\right)u. \tag{2.22}$$

Für die letzte Behauptung ist zu prüfen, ob bei einem Lipschitz-Gebiet

$$\gamma\left(u|_{\partial G}\right) = u|_{\partial G},\ u\in C^l(\overline{G})$$

erfüllt ist. Die Existenz eines totalen Erweiterungsoperators gewährleistet $E(u|_{\partial G}) = u|_{\partial G}$ (vgl. Bemerkung 2.2.5). Das Bild von $(\psi^j)^{-1}$ liegt in $W^{l,p}(U^j)\cap C(\overline{U^j})$ und liefert speziell $(\psi^j)^{-1}(\omega^j u|_{\partial G}) = (\psi^j)^{-1}(\omega^j u)|_{B_0}$. Bei diesem Argument wirkt $\tilde{\gamma}$ nach Definition als Identität, so dass mit $\psi^j((\psi^j)^{-1}(\omega^j u)|_{B_0}) = \omega^j u|_{\partial G}$ und $\sum_{j=1}^{r}\omega^j u|_{\partial G} = u|_{\partial G}$ alles gezeigt ist. □

Auch wenn auf Grund mangelnder Glattheit von u die Identität $\gamma u = u|_{\partial G}$ nicht gilt, so soll unter $u|_{\partial G}$ stets das Bild des Spuroperators verstanden werden. Ferner lehrt der Spursatz, dass die Regularität des Randes den maximal erreichbaren Differentiationsindex l des Sobolev-Raumes $W^{l,p}(\partial G)$ beschränkt.

Bemerkung 2.2.10 Die lokale Argumentation im Beweis des Spurlemmas sorgt dafür, dass für jede offene Überdeckung U^j

$$\|\gamma u\|_{l-1/p,p,U^j\cap\partial G} \leq C\|u\|_{l,p,G}$$

gilt. Ist also G ein Gebiet mit lokal höherer Glattheit als C^m $[C^{m,1}]$, so ist lokal, vermöge der Unabhängigkeit von der speziellen Wahl der offenen Überdeckungen, $u|_{\partial\tilde{G}} \in W^{l,p}(\partial\tilde{G})$ mit $l > m$, $[l > m+1]$ für $\tilde{G}\subset G$ gewährleistet.

2.3 Wichtige Gleichungen und Ungleichungen

Im weiteren Verlauf werden folgende Hilfsmittel benötigt:

Lemma 2.3.1 (verallgemeinerte *Youngsche Ungleichung*) *Für die Exponenten $p, q \in (1, \infty)$ mit $1/p + 1/q = 1$ und jedes $\epsilon > 0$ gilt:*

$$ab \leq \frac{1}{p}\epsilon a^p + \frac{1}{q}\frac{1}{\epsilon^{q/p}}b^q \quad \forall a, b \geq 0. \tag{2.23}$$

Beweis Z.B. [War99] mit $ab = (\epsilon^{1/p}a)(\epsilon^{-1/p}b)$. □

Lemma 2.3.2 (Ungleichung von *Poincaré-Friedrichs*) *Sei $1 \leq p < \infty$ und G ein beschränktes, zusammenhängendes Gebiet mit Lipschitz-Rand. Dann ist*

$$V = \{v \in W^{1,p}(G) : \int_G v\,dx = 0\}$$

ein abgeschlossener Unterraum von $W^{1,p}(G)$ und es existiert eine Konstante $C_{p,G} > 0$ derart, dass

$$C_{p,G}\|v\|_{1,p,G} \leq \|\nabla v\|_{0,p,G} \quad \forall v \in V. \tag{2.24}$$

Beweis [EG04, Theorem B.37, Lemma B.66]. □

Bemerkung 2.3.3 Der Raum V ist identisch mit dem Raum $\{v \in W^{1,p}(G) : v \neq \text{const}\}$, denn es bilden $V_1 = \{v \in W^{1,p}(G) : \int_G v\,dx = 0\}$ und $V_2 = \{v \in W^{1,p}(G) : v = \text{const}\}$ eine direkte Summe: Jedes $v \in L^p(G)$ lässt sich darstellen als $v = w + v - w$ mit $w = 1/|G| \int_G v\,dx \in V_2$ und $v - w \in V_1$, da $\int_G (v - w)\,dx = \int_G v\,dx - \int_G v\,dx = 0$.

Lemma 2.3.4 (*Deny und Lions*) *Sei $k \in \mathbb{N}_0$, $p \in [1, \infty]$ und $G \in \mathbb{R}^d$ ein Gebiet. Dann existiert ein $C_{DL} = C_{DL}(d, k, G) > 0$, derart, dass*

$$\inf_{p \in P_k(G)} \|v + p\|_{k+1,p,G} \leq C_{DL}|v|_{k+1,p,G} \quad \forall v \in W^{k+1,p}(G). \tag{2.25}$$

Beweis [Cia91, Theorem 14.1]. □

Bemerkung 2.3.5 Vgl. auch die Bemerkungen zur Abhängigkeit von C_{DL} in [Ape04, S. 61].

Lemma 2.3.6 (*Höldersche Ungleichung*) *Es seien $p, q, r \in [1, \infty]$ mit $1/r = 1/p + 1/q$. Dann ist $uv \in L^r(G)$ und mit*

$$\|uv\|_{0,r,G} \leq \|u\|_{0,p,G}\|v\|_{0,q,G} \tag{2.26}$$

für $u \in L^p(G), v \in L^q(G)$. Im Fall $u = v$ mit $r = 1, p, q = 2$ und $r, p, q = \infty$ sind beide Seiten gleich.

Beweis Für $p, q, r \in [1, \infty)$ siehe [War99]. Sei $r, p = 1$ und $q = \infty$. Damit ist $|v| \leq C$ fast überall für ein $C \in [0, \infty)$ und es gilt fast überall $|uv| \leq C|u|$. Sei N die entsprechende Nullmenge. Dann gilt

$$\|uv\|_{0,1,G} = \int_G |uv|\,dx = \int_{G\setminus N} |uv|\,dx \leq \int_{G\setminus N} C|u|\,dx = C\|u\|_{0,1,G}.$$

Da $\|v\|_{0,\infty,G}$ das Infimum aller dieser C ist, gilt auch $\|uv\|_{0,1,G} \leq \|v\|_{0,\infty,G}\|u\|_{0,1,G}$. Ferner gilt aufgrund der Monotonie von $t \mapsto t^m$, $t \in \mathbb{R}_+$ und $m \in \mathbb{N}$ für

$$\operatorname*{ess\,sup}_{x\in G}(|u(x)|) = |u(x^\sharp)|,\ x^\sharp \in G \setminus N$$

die Folgerung

$$|u(x)| \leq |u(x^\sharp)| \Rightarrow |u(x)|^m \leq |u(x^\sharp)|^m \quad \forall x \in G \setminus N. \tag{2.27}$$

Dieses mündet zusammen mit den Definitionen der verwendeten Normen in

$$\begin{aligned}\|u\|^m_{0,\infty,G} &= \left\{\operatorname*{ess\,sup}_{x\in G}(|u(x)|)\right\}^m = |u(x^\sharp)|^m \underset{(2.27)}{=} \operatorname*{ess\,sup}_{x\in G}(|u(x)|^m) = \operatorname*{ess\,sup}_{x\in G}(|u(x)^m|)\\ &= \|u^m\|_{0,\infty,G}.\end{aligned}$$

□

Lemma 2.3.7 (Interpolationsungleichung I) *Seien $p_0, p_1, p_\theta \in [1,\infty]$ und $0 \leq \theta \leq 1$ mit $\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$. Dann ist*

$$\forall u \in L^{p_0}(G) \cap L^{p_1}(G), \quad \|u\|_{0,p_\theta,G} \leq \|u\|^{1-\theta}_{0,p_0,G}\|u\|^\theta_{0,p_1,G}. \tag{2.28}$$

Beweis Setze in der Hölderschen Ungleichung (2.26) $f = u^\theta$, $g = u^{1-\theta}$:

$$\begin{aligned}\|u\|_{0,p_\theta,G} &\leq \|u^{1-\theta}\|_{0,\frac{p_0}{1-\theta},G}\|u^\theta\|_{0,\frac{p_1}{\theta},G}\\ &= \left(\int_G |u|^{p_0}\right)^{\frac{1-\theta}{p_0}}\left(\int_G |u|^{p_1}\,dx\right)^{\frac{\theta}{p_1}} = \|u\|^{1-\theta}_{0,p_0,G}\|u\|^\theta_{0,p_1,G}.\end{aligned}$$

□

Lemma 2.3.8 (Interpolationsungleichung II) *Sei $l_0, l_1 \in \mathbb{R}_+$, $l_0 \neq l_1$, $1 \leq p_0, p_1 \leq \infty$ und G ein beschränktes Gebiet mit Lipschitz-Rand. Für $0 < \theta < 1$ definiere $l_\theta = (1-\theta)l_0 + \theta l_1$, $\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$, $l_0, l_1 \in \mathbb{R}_+ \setminus \mathbb{N}_0$ bzw. $p_\theta = p_0 = p_1$, $l_0, l_1 \in \mathbb{N}_0$ und es gilt*

$$\forall u \in W^{l_0,p_0}(G) \cap W^{l_1,p_1}(G), \quad \|u\|_{l_\theta,p_\theta,G} \leq C\|u\|^{1-\theta}_{l_0,p_0,G}\|u\|^\theta_{l_1,p_1,G}. \tag{2.29}$$

Beweis Mit [Tri78, 1.3.3 (g)] und [BL76, Definition 6.2.2, Theorem 6.2.3, Theorem 6.2.4 und Theorem 6.4.5 (3),(4)] folgt die Behauptung für $G = \mathbb{R}^d$. Der Lipschitz-Rand garantiert die Existenz eines totalen Erweiterungsoperators (Lemma 2.2.7):

$$\begin{aligned}\|u\|_{l_\theta,p_\theta,G} &\underset{(2.17)}{=} \|Eu\|_{l_\theta,p_\theta,G} \leq \|Eu\|_{l_\theta,p_\theta,\mathbb{R}^d}\\ &\leq\ C\|Eu\|^{1-\theta}_{l_0,p_0,\mathbb{R}^d}\|Eu\|^\theta_{l_1,p_1,\mathbb{R}^d} \underset{(2.18)}{\leq} C\|u\|^{1-\theta}_{l_0,p_0,G}\|u\|^\theta_{l_1,p_1,G}.\end{aligned}$$

□

Korollar 2.3.9 *Sei $l_0, l_1 \in \mathbb{R}_+$, $l_0 \neq l_1$, $1 \leq p_0, p_1 \leq \infty$ und $G \subset \mathbb{R}^d$ ein beschränktes Gebiet mit Lipschitz-Rand. Für $0 < \theta < 1$ definiere $l_\theta = (1-\theta)l_0 + \theta l_1$, $\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$, $l_0, l_1 \in \mathbb{R}_+ \setminus \mathbb{N}_0$ bzw. $p_\theta = p_0 = p_1$, $l_0, l_1 \in \mathbb{N}_0$. Es folgt*

$$\forall u \in W^{l_0,p_0}(G) \cap W^{l_1,p_1}(G), \quad |u|_{l_\theta,p_\theta,G} \leq C|u|^{1-\theta}_{l_0,p_0,G}|u|^\theta_{l_1,p_1,G}. \tag{2.30}$$

Beweis Es genügt die Abschätzung für $l_0 = 0$ und $l_1 = 1$ zu zeigen, da der allgemeine Fall aus dem speziellen Resultat hervorgeht. Ist $u = \text{const}$, so ist die Ungleichung durch die Definition der Halbnormen trivial erfüllt. Sei also $u \neq \text{const}$. Die Anwendung der Ungleichung von Poincaré-Friedrichs (2.24) liefert mit (2.29)

$$\forall u \in W^{0,p_0}(G) \cap W^{1,p_1}(G), \quad \|u\|^p_{0,p,G} + |u|^p_{\theta,p_\theta,G} \leq C^p|u|^{p(1-\theta)}_{0,p_0,G}|u|^{p\theta}_{1,p_1,G}$$

die gewünschte Aussage. □

Lemma 2.3.10 (*Leibnizsche Produktregel*) *Für einen Multiindex α und zwei Funktionen f, g, deren Ableitungen bis zur Ordnung $|\alpha|$ existieren, gilt die Leibnizsche Produktregel*

$$\partial^\alpha (f \cdot g) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^\beta f \partial^{\alpha-\beta} g. \tag{2.31}$$

Beweis [Che01, Theorem 5.4.3]. □

Lemma 2.3.11 (*Gronwallsche Ungleichung*) *Sei $T \in \mathbb{R}_+ \cup \infty, t_0 \in [0,T), a, \alpha, \in L^\infty(t_0,T)$ und $\beta \in L^1(t_0,T), \beta(t) \geq 0$ fast überall in (t_0,T). Ist die Ungleichung*

$$a(t) \leq \alpha(t) + \int_{t_0}^t \beta(s) a(s)\, ds$$

für fast alle $t \in (t_0,T)$ erfüllt, so gilt für ebensolches t

$$a(t) \leq \alpha(t) + \int_{t_0}^t \exp\left(\int_s^t \beta(\tau)\, d\tau\right) \alpha(s)\beta(s)\, ds.$$

Ist α nicht fallend, so gilt auch

$$a(t) \leq \alpha(t) \exp\left(\int_{t_0}^t \beta(\tau)\, d\tau\right) \quad \forall t \in (t_0,T).$$

Beweis [Emm04] oder [QV94, Lemma 1.4.1]. □

Lemma 2.3.12 (*diskrete Gronwallsche Ungleichung*) *Ist γ_n eine nichtnegative Folge und erfüllt a_n die Bedingungen*

$$\begin{aligned} a_0 &\leq \alpha_0 \\ a_n &\leq \alpha_0 + \sum_{s=0}^{n-1} \beta_s + \sum_{s=0}^{n-1} \gamma_s a_s,\ n \geq 1, \end{aligned}$$

so gilt, falls $\alpha_0 \geq 0$ und $\beta_n \geq 0$ für $n \geq 0$

$$a_n \leq \left(\alpha_0 + \sum_{s=0}^{n-1} \beta_s\right) \exp\left(\sum_{s=0}^{n-1} \gamma_s\right),\ n \geq 1.$$

Beweis [QV94, Lemma 1.4.2]. □

3 Hyperbolische Erhaltungsgleichungen auf beschränkten Gebieten

Sei $Q_T = (0,T)\times\Omega \subset \mathbb{R}_+ \times \mathbb{R}^{d_x}, T > 0$ ein *Zeit-Raum-Zylinder* mit dem Rand $\Sigma_T = (0,T)\times \partial\Omega$ und es gelte $\partial\Omega = \Gamma_D = \Gamma_D^- \cup \Gamma_D^+$ mit $\Gamma_D^- \cap \Gamma_D^+ = \emptyset$. Man nennt

$$\Gamma_D^- = \{x \in \Gamma_D : {f_x}'(g_D)\cdot n < 0\}$$

den *Einflussrand* und Γ_D^+ den *Ausflussrand*. Mit n sei der äußere Normaleneinheitsvektor bezeichnet. Σ_D^+ und Σ_D^- werden analog definiert. Betrachtet wird für $u : \overline{Q}_T \to \mathbb{R}$ die Anfangs-Randwertaufgabe.

$$L(u) = \frac{\partial u}{\partial t} + \nabla \cdot f_x(u) = 0 \text{ in } Q_T, \tag{3.1}$$

$$„u = g_D \text{ auf } \Sigma_T“, \tag{3.2}$$

$$u(0,\cdot) = u_0 \text{ auf } \Omega, \tag{3.3}$$

mit der Flussdichte $f_x : \mathbb{R} \to \mathbb{R}^{d_x}$ und den Daten $g_D : \Sigma_T \to \mathbb{R}$ und $u_0 : \Omega \to \mathbb{R}$. Wird der Lösungsbegriff für (3.1) von den klassischen auf die schwachen Lösungen erweitert, so kann die Eindeutigkeit i.Allg. nicht mehr sichergestellt werden (vgl. u.a. ([War99, S. 247])). Als Ausweg wird ein regularisiertes parabolisches Hilfsproblem

$$\begin{aligned} -\epsilon\Delta u_\epsilon + L(u_\epsilon) &= 0 \text{ in } Q_T, \\ u_\epsilon &= g_{\epsilon D} \text{ auf } \Sigma_T, \\ u_\epsilon(0,\cdot) &= u_{\epsilon 0} \text{ auf } \Omega, \end{aligned} \tag{3.4}$$

für $u_\epsilon : \overline{Q}_T \to \mathbb{R}$ eingeführt, dass für $\epsilon \to 0$, $\epsilon > 0$ in das hyperbolische Problem (3.1) degenerieren soll (Viskositätsmethode). Für Existenz und Eindeutigkeit von (3.4) siehe z.B. [Fri64]. Damit ist für den Fall $\epsilon \to 0$ unter Beachtung der später genauer erläuterten Rand- und Anfangsbedingungen ein Kriterium gegeben, das dem Problem (3.1) eine eindeutige Lösung zuordnet. Die physikalische Interpretation der Viskositätsmethode, ist die Tatsache, dass ein reibungsfreier Prozess, als Grenzfall einer Folge von reibungsbehafteten Prozessen aufgefasst werden kann. Die beiden Formulierungen des zweiten Hauptsatzes der Thermodynamik

Alle Prozesse, bei denen Reibung auftritt sind irreversibel.

und

Alle natürlichen Prozesse sind irreversibel. Reversible Prozesse sind nur idealisierte Grenzfälle irreversibler Prozesse.

lassen die Viskositätsmethode als mathematisches Analogon des Hauptsatzes erscheinen (vgl. hierzu [AS97]). Für ein numerisches Verfahren mit $h \to 0 \Rightarrow \epsilon \to 0$ kann somit Eindeutigkeit erwartet werden, falls es in einem gewissen Sinn konform mit dem zweiten Hauptsatz der Thermodynamik ist.

3.1 Lösungstheorie in $BV(Q_T)$

Nachdem die Viskositätsmethode physikalisch motiviert wurde, folgt nun das mathematische Fundament für Lösungen $u \in BV(Q_T)$ von (3.1). Die genaue Formulierung der Anfangs- und Randbedingungen wird ebenfalls konkretisiert.

Lemma 3.1.1 *Für $u \in BV(Q_T)$ existieren die Spuroperatoren $\gamma_1 : BV(Q_T) \to L^\infty(\Omega)$ mit*

$$\operatorname*{ess\,lim}_{t \to 0+} \int_\Omega |u(t,x) - \gamma_1 u(0,x)|\, dx = 0$$

und $\gamma_2 : BV(Q_T) \to L^\infty(\Sigma_T)$ mit

$$\operatorname*{ess\,lim}_{s \to 0-} \int_{\Sigma_T} |u(r + sn(r)) - \gamma_2 u(r)|\, dr = 0.$$

Ferner gilt für $h \in C^1(\mathbb{R})$

$$\gamma_i(h(u)) = h(\gamma_i(u)), \quad i = 1, 2$$

fast überall auf Σ_T.

Beweis [BRN79, Lemma 1]. □

Annahme 3.1.2

1. $f_x \in C^2(\mathbb{R})^{d_x}$
2. ∇f_x ist global *Lipschitz-stetig*
3. $(u_{\epsilon 0}, g_{\epsilon D}) \in C^2(\overline{\Omega}) \times C^2(\Sigma_T)$.

Theorem 3.1.3 (Existenz der Entropie-Lösung) *Es gelte die Annahme* 3.1.2. *Die Lösungen $(u_\epsilon)_{\epsilon>0}$ des Problems (3.4) sind folgenkompakt in $L^1((0,T)\times\Omega)$ und für die Grenzfunktion gilt $u \in BV((0,T)\times\Omega)$ mit der* Randbedingung (3.3) *für $t = 0$.*

Beweis [BRN79, Theorem 1]. □

Damit das hyperbolische Problem (3.1) wohlgestellt ist, darf die *Dirichlet-Randbedingung* nur auf Σ_D^- gefordert werden: Sei $x \in \Sigma_D^-$ Startpunkt einer in Q_T verlaufenden Charakteristik, dann ist die Lösung u der homogenen Gleichung (3.1) auf der Charakteristik konstant, so dass eine Randbedingung auf Σ_D^+ i.Allg. einen Widerspruch hervorrufen würde. Aus diesem Grund kann für das Problem (3.4) die dazugehörige Randbedingung auf Σ_T für den Fall $\epsilon \to 0$ nicht mehr erfüllt werden. Falls eine schwächere Randbedingung für (3.1) auf Σ_T gefunden werden kann, die auch von der Grenzfunktion u für $\epsilon \to 0$ erfüllt wird, so ist die Viskositätsmethode wohldefiniert.

Es kann gezeigt werden (vgl. [BRN79, Theorem 2]), dass für die Grenzfunktion $\lim\limits_{\epsilon\to 0} u_\epsilon = u$ mit $u \in BV((0,T)\times\Omega)$ von (3.4) und alle nichtnegativen $\phi \in C^2((0,T)\times\overline{\Omega}), \phi \geq 0$ mit kompakten Träger

$$\begin{aligned} &\int_\Omega \int_0^T |u-k| \frac{\partial \phi}{\partial t} + \operatorname{sign}(u-k)\,(f_x(u) - f_x(k)) \cdot \nabla \phi \, dxdt \\ &- \int_{\Gamma_D} \int_0^T \operatorname{sign}(g_D - k)\,(f_x(\gamma_2 u) - f_x(k)) \cdot n\phi \, dsdt \geq 0,\ k \in \mathbb{R} \end{aligned} \tag{3.5}$$

gilt. Die Funktion $\operatorname{sign} : \mathbb{R} \to \mathbb{R}$ ist definiert durch

$$\operatorname{sign}(x) = \begin{cases} x/|x|, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Betrachte nun die spezielle Wahl $\phi = \Psi\rho_\delta$ für ein nichtnegatives $\Psi \in C^2((0,T))$ mit kompakten Support in $(0,T)$ und ein $\rho_\delta \in C^2(\overline{\Omega}),\ \delta > 0$ mit den Eigenschaften:

$$\begin{aligned} \rho_\delta &\equiv 1 \text{ auf } \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) \leq 1/2\,\delta\}, \\ \rho_\delta &\equiv 0 \text{ auf } \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) \geq \delta\}, \\ 0 \leq \rho_\delta &\leq 1 \text{ auf } \Omega, \\ \|\nabla\rho_\delta\|_{0,\infty,\Omega} &\leq \frac{c}{\delta} \text{ auf } \Omega, \end{aligned}$$

wobei c eine von δ unabhängige Konstante ist und

$$\operatorname{dist}(A,B) = \inf_{x\in A,\, y\in B} \|x - y\|$$

die Distanz zweier nichtleerer Mengen $A,\ B$ bezeichnet. Für $\delta \to 0$ entsteht aus (3.5) :

$$\begin{aligned} &\int_\Omega \int_0^T \operatorname{sign}(u-k)\,(f_x(u) - f_x(k)) \cdot \Psi\nabla\rho_\delta \, dxdt \\ &- \int_{\Gamma_D} \int_0^T \operatorname{sign}(g_D - k)\,(f_x(\gamma_2 u) - f_x(k)) \cdot n\Psi \, dsdt \geq 0 \end{aligned} \tag{3.6}$$

und damit nach partieller Integration

$$\int_{\Gamma_D} \int_0^T (\operatorname{sign}(\gamma_2 u - k) - \operatorname{sign}(g_D - k))\,(f_x(\gamma_2 u) - f_x(k)) \cdot n\Psi \, dsdt \geq 0. \tag{3.7}$$

Es ergibt sich also eine Randbedingung auf Σ_T, die von der Grenzfunktion u für $(t,x) \in \Sigma_T$ erfüllt wird:

$$(\operatorname{sign}(\gamma_2 u(t,x) - k) - \operatorname{sign}(g_D(t,x) - k))\,(f_x(\gamma_2 u(t,x)) - f_x(k)) \cdot n \geq 0,\ \forall k \in \mathbb{R}. \tag{3.8}$$

Für $k \notin \mathcal{I}(\gamma_2 u, g_D)$, $\mathcal{I}[a,b] = [\min(a,b), \max(a,b)]$ folgt durch einfaches Nachrechnen, dass (3.8) automatisch erfüllt ist. Somit ist für alle $k \in \mathcal{I}(\gamma_2 u, g_D)$

$$(\operatorname{sign}(\gamma_2 u(t,x) - k) - \operatorname{sign}(g_D(t,x) - k))\,(f_x(\gamma_2 u(t,x)) - f_x(k)) \cdot n \geq 0, \tag{3.9}$$

eine äquivalente Aussage. Für $\mathcal{I}(\gamma_2 u, g_D)$ ergibt sich eine weitere Äquivalenz durch

$$(\operatorname{sign}(\gamma_2 u(t,x) - g_D(t,x)))\,(f_x(\gamma_2 u(t,x)) - f_x(k)) \cdot n \geq 0,\ \forall k \in \mathcal{I}(\gamma_2 u, g_D). \tag{3.10}$$

Andererseits sind die äquivalenten Bedingungen für eine lineare Funktion f mit $\gamma_2 u = g_D$ im Einflussrand immer erfüllt. Denn für $f_x{}' \cdot n \geq 0$, $n = n(x_0)$, x_0 fest gewählt, folgt für jedes $u \in BV((0,T)\times\Omega)$ aus $f_x{}' \cdot n \geq 0$ lokal für $f_x(\gamma_2 u) \cdot n \geq f_x(k) \cdot n\,[f_x(\gamma_2 u) \cdot n < f_x(k) \cdot n]$

die Beziehung $\gamma_2 u \geq k\,[\gamma_2 u < k]$ und somit gilt ebenfalls (3.8), denn aus dem Mittelwertsatz folgt

$$\begin{aligned}(f_x(\gamma_2 u) - f_x(k)) \cdot n &= f'(\xi) \cdot n\,(\gamma_2 u - k) \\ &= f'(\xi) \cdot n\,(\gamma_2 u - k),\ \xi \in \mathcal{I}(\gamma_2 u, k).\end{aligned} \tag{3.11}$$

Für eine nichtlineare Funktion f kann im Ausflussrand $({f_x}'(g_D) \cdot n \geq 0)$ i.Allg. nicht mehr ${f_x}'(\xi) \cdot n \geq 0$ erwartet werden.

Mit der nachfolgenden Definition lässt sich die Eindeutigkeit der Entropie-Lösung sicherstellen.

Definition 3.1.4 Eine Funktion $u \in BV((0,T) \times \Omega)$ ist eine Lösung des Problems (3.1), (3.8), (3.3), wenn für alle $k \in \mathbb{R}$ und alle nichtnegativen Testfunktionen $\phi \in C^2((0,T) \times \overline{\Omega})$ mit kompakten Support in $(0,T) \times \overline{\Omega}$ die Bedingung (3.5) und die Anfangswertbedingung (3.3) fast überall auf Ω erfüllt.

Theorem 3.1.5 (Eindeutigkeit der Entropie-Lösung) *Es gelte die Annahme* 3.1.2. *Dann ist die Lösung des Problems* (3.1), (3.8), (3.3) *eindeutig und identisch mit der Lösung der Viskositätsmethode.*

Beweis [BRN79, Theorem 2]. □

Damit ist das Ausgangsproblem unter der Annahme 3.1.2 und für Funktionen $u \in BV(Q_T)$ wohlgestellt und kann wie folgt definiert werden.

$$\nabla \cdot (f(u)) = 0 \text{ in } Q_T, \tag{3.12}$$

$$(\operatorname{sign}(\gamma_2 u - k) - \operatorname{sign}(g_D - k))\,(f(\gamma_2 u) - f(k)) \cdot n \geq 0 \text{ auf } \Sigma_T, \tag{3.13}$$

$$u(0,\cdot) = u_0 \text{ auf } \Omega, \tag{3.14}$$

wobei $f : \mathbb{R} \to \mathbb{R}^d$, $d = d_x + 1$ mit $f_0(u) = u$ und $\nabla\cdot$ die Divergenz bzgl. $x = (x_0, x_1, \ldots, x_{d_x})$ mit $x_0 = t$ bezeichnet.

3.2 Lösungstheorie in $L^\infty(Q_T)$

Der in der Einleitung dieses Kapitels angedeutete Weg zu einer eindeutigen schwachen Lösung von (3.1) zu gelangen, ist motiviert durch physikalische Betrachtungen der Entropie. Hierzu seien $\eta \in C^1(\mathbb{R})$ konvex und $q \in C^1(\mathbb{R})^{d_x}$ Funktionen, die für glatte Lösungen $u \in C^1(\mathbb{R}^{d_x})$ von (3.1) automatisch eine weitere Erhaltungsgleichung

$$\frac{\partial}{\partial t}\eta(u) + \nabla \cdot q(u) = 0 \tag{3.15}$$

erfüllen. Wird (3.1) mit $\eta'(u)$ multipliziert, so muss die *Kompatibilitätsbedingung*

$$\eta'(u)\nabla f_x(u) = \nabla q(u), \quad \forall u \in \mathbb{R} \tag{3.16}$$

erfüllt sein, damit (3.15) gilt. Man nennt η, q in diesem Fall *Entropie-Flux-Paar*. Für skalare Erhaltungsgleichungen erzeugt jede konvexe Funktion $\eta \in C^1(\mathbb{R})$ für fest gewähltes $k \in \mathbb{R}$ vermöge

$$q_j(u) = \int_k^u \eta'(r) f_j'(r)\, dr, \quad 1 \le j \le d_x,$$

ein Entropie-Flux-Paar.

Ist $u \notin C^1(\mathbb{R}^{d_x})$ nicht glatt genug, so kann mit der Viskositätsmethode (vgl. [Mal+96, S. 61]) und (3.16) folgende Definition gerechtfertigt werden:

Definition 3.2.1 Eine schwache Lösung von (3.1) mit $\Omega = \mathbb{R}^{d_x}$ wird *Entropie-Lösung* genannt, wenn für alle Entropie-Flux-Paare die *Entropie-Ungleichung*

$$\frac{\partial}{\partial t}\eta(u) + \nabla \cdot q(u) \le 0 \tag{3.17}$$

im schwachen Sinn erfüllt ist.

Im Gegensatz zu η ist die physikalische Entropie eine konkave Funktion, die bei Stößen, d. h. zulässigen Unstetigkeiten, zunimmt.

Die Formulierung der schwachen Rand- und Anfangswertbedingungen ist mit dem Problem verbunden, dass Funktionen aus $L^\infty(Q_T)$ nach dem Spurtheorem 2.2.9 i.Allg. keine Randwerte besitzen. Für Funktionen aus $BV(Q_T)$ konnte in diesem Zusammenhang auf Lemma 3.1.1 zurückgegriffen werden. Werden die Eigenschaften des Spuroperators von Lemma 3.1.1 jedoch in die schwachen Rand- und Anfangswertbedingungen integriert, so kann eine wohlgestellte, schwache Formulierung des Rand- und Anfangswertproblems gefunden werden, die eine eindeutige Lösung erlaubt:

Definition 3.2.2 Sei $u_0 \in L^\infty(\Omega)$, $g_D \in L^\infty(\Sigma_T)$ und $f_x \in C^1(\mathbb{R})^{d_x}$. u wird schwache Lösung von (3.1), (3.8) und (3.3) genau dann genannt, wenn

$$u \in L^\infty(Q_T), \tag{3.18}$$

und die folgenden Bedingungen erfüllt sind:

1. Die *Entropie-Bedingung* im Sinn von

$$\int_{Q_T} \eta(u)\frac{\partial \varphi}{\partial t} + \sum_{i=1}^{d_x} q_i(u)\frac{\partial \varphi}{\partial x_i}\, d(t,x) \ge 0 \tag{3.19}$$

   für alle $\varphi \in C_0^\infty(Q_T)$, $\varphi \ge 0$ und alle Entropie-Flux-Paare $(\eta,\, q)$;

2. die Randbedingung $g_D \in L^\infty(\Sigma_T)$ für $k \in \mathcal{I}[u^\tau(r), g_D(r)]$, und fast alle $r \in \Sigma_T$ im Sinn von

$$(\operatorname{sign}(u^\tau(r) - k) - \operatorname{sign}(g_D(r) - k))\,(f_x(u^\tau(r)) - f_x(k)) \cdot n(r) \ge 0, \tag{3.20}$$

   falls ein $u^\tau \in L^\infty(\Sigma_T)$ existiert mit

$$\operatorname*{ess\,lim}_{s \to 0-} \int_{\Sigma_T} |u(r + sn(r)) - u^\tau(r)|\, dr = 0; \tag{3.21}$$

3. die Anfangswertbedingung $u_0 \in L^\infty(\Omega)$ im Sinn von

$$\operatorname*{ess\,lim}_{t\to 0+} \int_\Omega |u(t,x) - u_0(x)|\, dx = 0. \tag{3.22}$$

Bemerkung 3.2.3 Diese Definition entspricht [Mal+96, Definition 7.2], wobei aber die dortige Randbedingung mit einer nach [Mal+96, Definition 7.24] äquivalenten Bedingung ausgetauscht wurde. Vgl. dazu auch die äquivalenten Formulierungen aus dem letzten Abschnitt. Die Bedingung (3.19) ergibt für $\eta = \pm Id$ und $\varphi \in C_0^\infty(Q_T)$ gerade die schwache Formulierung

$$\int_{Q_T} u\frac{\partial \varphi}{\partial t} + \sum_{i=1}^{d_x} f_i(u)\frac{\partial \varphi}{\partial x_i}\, d(t,x) = 0, \quad \forall \varphi \in C_0^\infty(Q_T). \tag{3.23}$$

(vgl. [Mal+96, Remark 7.7]).

Mit dieser Definition kann die Existenz und Eindeutigkeit der Entropie-Lösung gewährleistet werden.

Annahme 3.2.4

1. $f_x \in C^2(\mathbb{R})^{d_x}$
2. $(u_{\epsilon 0}, u_{\epsilon D}) \in L^\infty(\Omega) \times L^\infty(\Sigma_T)$.

Theorem 3.2.5 (Existenz der Entropie-Lösung) *Es gelte* 3.2.4. *Für $\epsilon > 0$ bezeichne u_ϵ eine Lösung von* (3.4). *$g_{\epsilon D}$, $u_{\epsilon 0}$ seien gleichmäßig beschränkt bzgl. der L^∞-Norm und konvergieren:*

$$\lim_{\epsilon\to 0+} g_{\epsilon D} = g_D \text{ in } L^1(\Sigma_T)$$
$$\lim_{\epsilon\to 0+} u_{\epsilon 0} = u_0 \text{ in } L^1(\Omega),$$

wobei $g_D \in L^\infty(\Sigma_T)$ und $u_0 \in L^\infty(\Omega)$. Dann ist $(u_\epsilon)_\epsilon$ gleichmäßig beschränkt und konvergiert in $C^0([0,T], L^1(\Omega))$ gegen die Lösung $u \in L^\infty(Q_T)$ von (3.19) *mit der Randbedingung* (3.20) *und den Anfangsbedingungen von* (3.22).

Beweis [Mal+96, Theorem 8.20]. □

Bemerkung 3.2.6 Theorem $3.2.5$ zeigt die Notwendigkeit von $f_x \in C^2(\mathbb{R})^{d_x}$ für die Existenz einer Entropie-Lösung, während für die Eindeutigkeit $f_x \in C^1(\mathbb{R})^{d_x}$ ausreichend ist:

Annahme 3.2.7

1. $f_x \in C^1(\mathbb{R})^{d_x}$
2. $(u_i, (g_D)_i, (u_0)_i) \in L^\infty(Q_T) \times L^\infty(\Sigma_T) \times L^\infty(\Omega)$, $i = 1, 2$.

Theorem 3.2.8 (Eindeutigkeit der Entropie-Lösung) *Sei Annahme 3.2.7 erfüllt und seien u_1, u_2 schwache Lösungen, die* (3.19)*-*(3.22) *erfüllen. Dann gilt für*

$$F(z,k) = \operatorname{sign}(z-k)(f_x(z) - f_x(k))$$

und alle $\beta \in C_0^\infty((-\infty, T) \times \mathbb{R}^{d_x})$

$$
\begin{aligned}
&-\int_0^T \int_\Omega |u_1 - u_2| \frac{\partial \beta}{\partial t} + \sum_{i=1}^{d_x} F_i(u_1, u_2) \frac{\partial \beta}{\partial x_i} \, dx \, dt \\
&\leq \int_\Omega |(u_0)_1 - (u_0)_2| \beta(0) \, dx + \int_{\Sigma_T} \operatorname{diam}(f_x \cdot n, \mathcal{I}[(g_D)_1, (g_D)_2]) \beta \, dr
\end{aligned} \tag{3.24}
$$

mit

$$
\begin{aligned}
&\operatorname{diam}(f_x \cdot n(r)), \mathcal{I}[(g_d)_1(r), (g_d)_2(r)] \\
&= \sup \left\{ |f_x(z_1) \cdot n(r) - f_x(z_2) \cdot n(r)| \, : z_1, z_2 \in \mathcal{I}[(g_D)_1, (g_D)_2] \right\}.
\end{aligned}
$$

Ferner existiert für fast alle $t \in (0,T)$ eine Lipschitz-Konstante $\mathcal{L}_{[f_x]}$ *bzgl. einer Kugel K mit dem Radius* $\max\{\|u_i\|_{0,\infty,K}, \|(g_D)_i\|_{0,\infty,K}, \|(u_0)_i\|_{0,\infty,K}\}$ *mit*

$$
\int_\Omega |u_1(t) - u_2(t)| \, dx \leq \int_\Omega |(u_0)_1 - (u_0)_2| \, dx + \mathcal{L}_{[f_x]} \int_0^t \int_{\Gamma_D} |(g_D)_1 - (g_D)_2| \, dr \, ds. \tag{3.25}
$$

Definition 3.2.9 Das Paar

$$
(|u-k|, \operatorname{sign}(u-k)(f_x(u) - f_x(k)))
$$

wird *Kružkovsches Entropie-Paar* und

$$
\left((u-k)^\pm, \operatorname{sign}^\pm(u-k)(f_x(u) - f_x(k))\right)
$$

wird *Semi-Kružkovsches Entropie-Paar* genannt. Dabei werden $x^\pm = \operatorname{sign}^\pm(x)x$,

$$
\operatorname{sign}^+(x) = \begin{cases} 1, & x > 0, \\ 0, & sonst \end{cases} \quad \text{und} \quad \operatorname{sign}^-(x) = \begin{cases} -1, & x < 0, \\ 0, & sonst \end{cases}
$$

definiert.

Annahme 3.2.10

1. $f_x \in C^2(\mathbb{R})^{d_x}$
2. f_x ist global Lipschitz-stetig
3. $(u_{\epsilon 0}, u_{\epsilon D}) \in L^\infty(\Omega) \times L^\infty(\Sigma_T)$.

Lemma 3.2.11 *Es gelte* 3.2.10. *Ist zusätzlich $u \in L^\infty(Q_T)$, dann ist die Formulierung* (3.19), (3.20) *und* (3.22) *äquivalent mit der Bedingung*

$$
\begin{aligned}
&\int_{Q_T} (u-k)^\pm \frac{\partial \varphi}{\partial t} + (\operatorname{sign}^\pm(u-k)(f_x(u) - f_x(k))) \cdot \nabla \varphi \, dx \, dt \\
&+ \int_\Omega (u_0 - k)^\pm \varphi(0,x) \, dx \\
&+ \mathcal{L}_{[f_x]} \int_{\Sigma_T} (g_D - k)^\pm \varphi \, ds \, dt \geq 0 \quad \forall \varphi \in C_0^\infty((-\infty, T) \times \mathbb{R}^{d_x}),
\end{aligned} \tag{3.26}
$$

mit $\varphi \geq 0$, $k \in \mathbb{R}$ und der Lipschitz-Konstante $\mathcal{L}_{[f_x]}$ bzgl. einer Kugel K mit dem Radius $\max\{\|u\|_{0,\infty,K}, \|g_D\|_{0,\infty,K}, \|u_0\|_{0,\infty,K}\}$.

Beweis Mit $H(u,k) = (u-k)^{\pm}$ und $Q(u,k) = \operatorname{sign}^{\pm}(u-k)(f_x(u) - f_x(k))$ folgt die Behauptung aus [Mal+96, Definition 7.1, Theorem 7.31]. □

Wird das Semi-Kružkovsche Entropie-Paar wie im Fall $\Omega = \mathbb{R}^{d_x}$ (vgl. (3.5)) durch das Kružkovsche Entropie-Paar ersetzt, so zeigt ein Beispiel aus [Vov02, S. 5], dass keine Eindeutigkeit der schwachen Lösung mehr gewährleistet werden kann.

4 Grundlagen zur Finite-Elemente-Approximation

Die behandelten Themen dieses Kapitels sind die Eckpfeiler für die Analysis der Discontinuous-Galerkin-Methode. Sie stellen einerseits Ergebnisse aus Gebieten bereit, die für jedes Studium von Finite-Elemente-Methoden wichtig sind, wie z.B. die stückweise Polynomapproximation in Sobolev-Räumen, Quadraturformeln und inverse Ungleichungen, andererseits aber auch Definitionen und Aussagen, die speziell die technischen Schwierigkeiten behandeln, die bei der $L^\infty(L^\infty)$-Abschätzung durch die Projektion innerhalb des Shock-capturing Terms verursacht werden, wie z.B. das Lumping oder den Versuch eine hierarchische Basis über Knotenpunkte zu gewährleisten. Wenn möglich, wird dabei neben der Abhängigkeit der Gitterweite auch der Polynomgrad explizit mit aufgeführt.

4.1 Geometrie des Zeit-Raum Gebietes

Es seien $Q_{n,n+1} = (t_n, t_{n+1}) \times \Omega$, $Q_n = \{t_n\} \times \Omega$ und $Q_{n+1} = \{t_{n+1}\} \times \Omega$ für die Zeitpunkte $0 = t_0 < t_1 < \cdots < t_N, N \in \mathbb{N}$ eine *Partitionierung* des Zeit-Raum Zylinders Q_T. $\Sigma^{\pm}_{n,n+1} = (t_n, t_{n+1}) \times \Gamma^{\pm}_D$, $\Sigma^{\pm}_n = \{t_n\} \times \Gamma^{\pm}_D$ und $\Sigma^{\pm}_{n+1} = \{t_{n+1}\} \times \Gamma^{\pm}_D$ bezeichnen den Einflussrand der Partition. $\mathcal{T}^n_h$ ist eine Zeit-Raum Partitionierung im Sinne von [EG04, Definition 1.49] von $Q_{n,n+1}$ in Tensorproduktelemente T mit dem Durchmesser $h_T = \max\limits_{x,y \in T} \|x - y\|_{l^2}$ und dessen Maximum $h = \max\{h_T : T \in \mathcal{T}^n_h,\, n = 1, \ldots, N\}$. Ferner ist $\mathcal{T}_h = \bigcup_{n \geq 0} \mathcal{T}^n_h$. $R_{n,n+1}$, R_{n+1} und R_n bezeichnet alle Kanten, die in $Q_{n,n+1}$, Q_{n+1} bzw. Q_n enthalten sind. Analog kennzeichnet $\Lambda^{\pm}_{n,n+1}$, $\Lambda^{\pm}_{n+1}$ und $\Lambda^{\pm}_n$ die Kanten in $\Sigma^{\pm}_{n,n+1}$, $\Sigma^{\pm}_{n+1}$ und $\Sigma^{\pm}_n$. Die inneren Kanten werden über $R^i_{n,n+1} = R_{n,n+1} \setminus \Lambda_{n,n+1}$ definiert.

In dieser Arbeit werden die üblichen Definitionen von *Finiten-Elementen* $\{T, P, \Sigma\}$ benutzt, siehe hierzu z.B. [EG04, Definition 1.23].

Der verwendete Finite-Elemente-Raum wird nun mit Hilfe von

$$W^n_h = \left\{ w \in L^2(Q_{n,n+1}) : w|_T \in P(T) \quad \forall T \in \mathcal{T}^n_h \right\} \tag{4.1}$$

definiert durch $W_h = \prod_{n \geq 0} W^n_h$ und ist ein Unterraum von

$$W^{l,p}(Q_T, \mathcal{T}_h) = \prod_{n \geq 0} \left\{ w \in L^2(Q_{n,n+1}) : w|_T \in W^{l,p}(T) \quad \forall T \in \mathcal{T}^n_h \right\}, \tag{4.2}$$

da an $P(T)$ die Forderung $P(T) \subset W^{l,\infty}(T)$ gestellt wird.

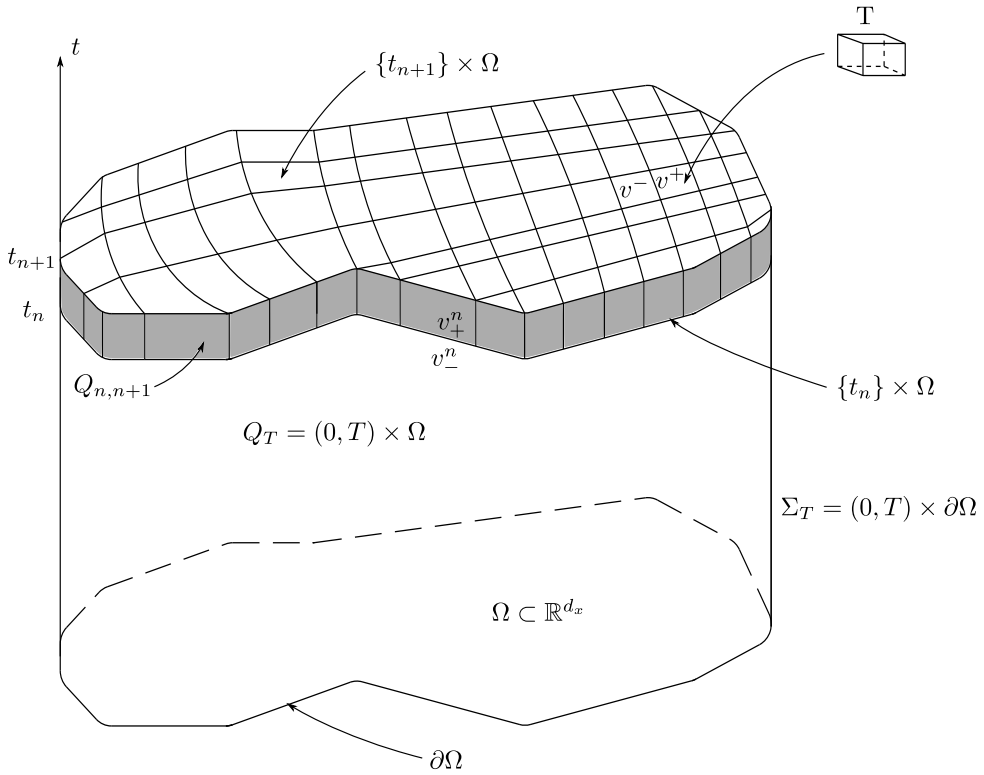


Abbildung 4.1: Geometrie des Zeit-Raum Gebietes

Das erwähnte Finite-Element $\{T, P, \Sigma\}$ wird aus einem Finiten-Referenzelement $\{\hat{T}, \hat{P}, \hat{\Sigma}\}$ erzeugt. Mit Hilfe eines C^1-*Diffeomorphismus* $F_T : \hat{T} \to T$ erfolgt die Transformation eines Referenzelements $\hat{T}$ auf das Element T.

Handelt es sich bei $\{\hat{T}, \hat{P}, \hat{\Sigma}\}$ um ein *Lagrangesches Finites-Element* (vgl. [EG04, Definition 1.27]) mit der Knotenmenge $\mathcal{N} = \left\{x_1, \ldots, x_{n_{\text{dof}}^k}\right\}$, $n_{\text{dof}}^k = \#\hat{\Sigma}$ und den Linearformen

$$\hat{\sigma}_i(\hat{v}) = \hat{v}(x_i),\ 1 \leq i \leq n_{\text{dof}}^k,\ \forall \hat{v} \in \hat{P},$$

so können die Definitionen

$$P(T) = \hat{P}(F_T^{-1}(T))$$

und

$$\sigma_i(v(x)) = \hat{\sigma}_i(v(F_T(\hat{x}))) = \hat{\sigma}_i(\hat{v}(\hat{x})),\ x = F_T(\hat{x}),\ 1 \leq i \leq n_{\text{dof}}^k$$

verwendet werden.

Die dazugehörige nodale Basis bestehend aus den Formfunktionen $\{\hat{\varphi}_1, \ldots, \hat{\varphi}_{n_{\text{dof}}^k}\}$ mit

$$\hat{\varphi}_i(x_j) = \delta_{ij},\ 1 \leq i, j \leq n_{\text{dof}}^k$$

werden entsprechend transformiert.

Die Partitionierung heißt *affin*, wenn

$$F_T : \hat{T} \ni \hat{x} \mapsto J_T\hat{x} + b_T \in T$$

mit $J_T \in \mathbb{R}^{d,d}$, $\det(J_T) \neq 0$, $b_T \in \mathbb{R}^d$ ist. Darüber hinaus soll $\mathcal{T}_h^n$ folgende Eigenschaften besitzen:

Definition 4.1.1 (lokal quasiuniform, shape regular) Eine Familie von affinen Partitionierungen $\{\mathcal{T}_h^n\}_{h>0}$ heißt *lokal quasiuniform*, wenn ein σ_0 derart existiert, dass gilt:

$$\forall h,\, \forall T \in \mathcal{T}_h^n, \quad \sigma_T = \frac{h_T}{\rho_T} \leq \sigma_0. \tag{4.3}$$

Hierbei bezeichnet ρ_T den Durchmesser der größten in T enthaltenen Kugel.

Definition 4.1.2 (quasiuniform) Eine Familie von Partitionierungen $\{\mathcal{T}_h^n\}_{h>0}$ heißt genau dann *quasiuniform*, wenn sie lokal quasiuniform ist und ein $C_{qu} > 0$ existiert mit

$$\forall h,\, \forall T \in \mathcal{T}_h^n, \quad h_T \geq C_{qu} h. \tag{4.4}$$

4.2 Polynomapproximation mit Tensorprodukten

Bei den später betrachteten Ungleichungen werden speziell die Eigenschaften der Tensorprodukt Finiten-Elemente genutzt. Das Referenzelement ist hier $\hat{T} = I^d$, $I = (-1, 1)$ und der Vektorraum wird definiert als $\hat{P} = \mathbb{Q}_k(\hat{T})$ mit

$$\mathbb{Q}_k(x) = \operatorname*{span}_{\alpha \in \mathbb{N}_0^d,\, \|\alpha\|_{l^\infty} \leq k} \{x^\alpha\},\, x \in T,\, k \in \mathbb{N}_0.$$

Die Formfunktion $\hat{\varphi}_\alpha \in \{\hat{\varphi}_1, \ldots, \hat{\varphi}_{n_{\text{dof}}^k}\}, \alpha \in \mathbb{N}_0^d$ lässt sich hierbei als Produkt der eindimensionalen *Lagrange-Polynome* schreiben: Für $x \in \mathbb{R}$ sind $\{\hat{\varphi}_0^k, \hat{\varphi}_1^k, \ldots, \hat{\varphi}_k^k\}$ die entsprechenden Polynome vom Grad k. Dadurch gilt für den Multiindex α mit $\|\alpha\|_{l^\infty} \leq k$ die Darstellung:

$$\hat{\varphi}_\alpha(\hat{x}) = \hat{\varphi}_{\alpha_0}^k(\hat{x}_0)\hat{\varphi}_{\alpha_1}^k(\hat{x}_1) \cdots \hat{\varphi}_{\alpha_{d_x}}^k(\hat{x}_{d_x}).$$

Die Transformation zwischen dem Referenzelement $\hat{T}$ und T geschieht in entsprechenden Abschätzungen mit dem folgenden Lemma, dessen Gestalt im Wesentlichen auf die affine Struktur der Transformation zurückzuführen ist.

Lemma 4.2.1 *Für $l \geq 0$ und $1 \leq p \leq \infty$, $0 = 1/\infty$ existiert ein $C_{l,d} \geq 1$ derart, dass für $T \in \mathcal{T}_h^n$, $\mathcal{T}_h^n$ affin und $w \in W^{l,p}(T)$, $\hat{w} = w \circ F_T$*

$$|\hat{w}|_{l,p,\hat{T}} \leq C_{l,d} \|J_T\|_{l^2}^l |\det(J_T)|^{-1/p} |w|_{l,p,T}, \tag{4.5}$$

$$|w|_{l,p,T} \leq C_{l,d} \|J_T^{-1}\|_{l^2}^l |\det(J_T)|^{1/p} |\hat{w}|_{l,p,\hat{T}}, \tag{4.6}$$

wobei $C_{l,d}$ nur von l und d abhängig ist. Speziell gilt $C_{0,d} = 1$.

Beweis Faà di Bruno's Formel ergibt für $\hat{w} = w \circ F_T$, $|\alpha| = l$, $P = \mathbb{N}_0^d \setminus \{0\}$ und

$$\mathcal{A} = \left\{ a : P \to \mathbb{N}_0^d : \sum_{\gamma \in P} a(\gamma) = \beta \text{ und } \sum_{\gamma \in P} |a(\gamma)|\gamma = \alpha \right\}$$

die Darstellung

$$\begin{aligned}\hat{\partial}^\alpha(w\circ F_T)(\hat{x}) &= \alpha!\sum_{|\beta|\le|\alpha|}(\partial^\beta w)(F_T(\hat{x}))\sum_{a\in\mathcal{A}}\prod_{\gamma\in P}\frac{1}{a(\gamma)!}\left[\frac{(\hat{\partial}^\gamma F_T)(\hat{x})}{\gamma!}\right]^{a(\gamma)}\\ &= \alpha!\sum_{|\beta|\le|\alpha|}(\partial^\beta w)(F_T(\hat{x}))\sum_{a\in\mathcal{A}}\prod_{j=1}^{d}\frac{1}{a(e_j)!}\left[(\hat{\partial}_j F_T)\right]^{a(e_j)}\\ &= \alpha!\sum_{|\beta|=l}(\partial^\beta w)(F_T(\hat{x}))\sum_{a\in\mathcal{A}}\prod_{j=1}^{d}\frac{1}{a(e_j)!}\left[(\hat{\partial}_j F_T)\right]^{a(e_j)},\end{aligned}$$

denn die Bedingungen $l=|\alpha|=|\sum_{\gamma\in\mathbb{N}^m}|a(\gamma)|\gamma|=|\sum_{j=1}^d|a(e_j)|e_j|=\sum_{j=1}^d|a(e_j)|$ und $|\beta|=|\sum_{j=1}^d a(e_j)|=\sum_{j=1}^d|a(e_j)|$ ergeben unter Vernachlässigung von leeren Summen die Summierungsbedingung $|\alpha|=|\beta|=l$. Für den Betrag der linken Seite ergeben sich die Abschätzungen

$$\begin{aligned}|\hat{\partial}^\alpha(w\circ F_T)(\hat{x})| &\le \alpha!\sum_{|\beta|=l}\left|(\partial^\beta w)(F_T(\hat{x}))\right|\sum_{a\in\mathcal{A}}\prod_{j=1}^{d}\frac{1}{a(e_j)!}\left|(\hat{\partial}_j F_T)^{a(e_j)}\right|\\ &\le \alpha!\sum_{|\beta|=l}\left|(\partial^\beta w)(F_T(\hat{x}))\right|\sum_{a\in\mathcal{A}}\prod_{j=1}^{d}\frac{1}{a(e_j)!}\left\|\hat{\partial}_j F_T\right\|_{l^\infty}^{|a(e_j)|}\\ &\le \alpha!\sum_{|\beta|=l}\left|(\partial^\beta w)(F_T(\hat{x}))\right|\sum_{a\in\mathcal{A}}\prod_{j=1}^{d}\frac{1}{a(e_j)!}\left\|J_T\right\|_{\max}^{l}\\ &= C_{\alpha,d}\sum_{|\beta|=l}\left|(\partial^\beta w)(F_T(\hat{x}))\right|\left\|J_T\right\|_{\max}^{l}\\ &\le C_{\alpha,d}\sum_{|\beta|=l}\left|(\partial^\beta w)(F_T(\hat{x}))\right|\left\|J_T\right\|_{l^2}^{l},\end{aligned}$$

so dass für das Referenzelement bei $p<\infty$ und $\sum_{|\beta|=l}1=\binom{d+l-1}{l}$ folgt

$$\|\hat{\partial}^\alpha\hat{w}\|_{0,p,\hat{T}}^p \underset{(2.6)}{\le} \binom{d+l-1}{l}^{(p-1)} C_{\alpha,d}^p\|J_T\|_{l^2}^{lp}\sum_{|\beta|=l}\|\partial^\beta w\circ F_T\|_{0,p,\hat{T}}^p$$

und nach Anwendung der Substitutionsregel

$$\|\hat{\partial}^\alpha\hat{w}\|_{0,p,\hat{T}}^p \le \binom{d+l-1}{l}^{(p-1)} C_{\alpha,d}^p\|J_T\|_{l^2}^{lp}|\det(J_T)|^{-1}|w|_{l,p,T}^p.$$

Aufsummierung über α liefert

$$\begin{aligned}|\hat{w}|_{l,p,\hat{T}}^p &= \sum_{|\alpha|=l}\|\hat{\partial}^\alpha\hat{w}\|_{0,p,\hat{T}}^p\\ &\le \sum_{|\alpha|=l}\binom{d+l-1}{l}^{(p-1)} C_{\alpha,d}^p\|J_T\|_{l^2}^{lp}|\det(J_T)|^{-1}|w|_{l,p,T}^p\\ &\le \left(\max_{|\alpha|=l}(C_{\alpha,d})\right)^p\binom{d+l-1}{l}^{p}\|J_T\|_{l^2}^{lp}|\det(J_T)|^{-1}|w|_{l,p,T}^p,\end{aligned}$$

mit $\binom{d+l-1}{l}\max_{|\alpha|=l}(C_{\alpha,d}) = C_{l,d}$ die Behauptung. Für $p = \infty$ liefert

$$\begin{aligned}
|\hat{w}|_{l,\infty,\hat{T}} &= \max_{|\alpha|=l}\|\hat{\partial}^{\alpha}\hat{w}\|_{0,\infty,\hat{T}} \\
&\leq \max_{|\alpha|=l}(C_{\alpha,d})\,\|J_T\|_{l^2}^{l}\|\sum_{|\beta|=l}\left|\left(\partial^{\beta}w\right)(F_T(\hat{x}))\right|\,\|_{0,\infty,\hat{T}} \\
&= \max_{|\alpha|=l}(C_{\alpha,d})\,\|J_T\|_{l^2}^{l}\|\sum_{|\beta|=l}\left|\left(\partial^{\beta}w\right)(x)\right|\,\|_{0,\infty,T} \\
&\leq \max_{|\alpha|=l}(C_{\alpha,d})\,\|J_T\|_{l^2}^{l}\binom{d+l-1}{l}\|\max_{|\beta|=l}\left|\partial^{\beta}w\right|\,\|_{0,\infty,T} \\
&= C_{l,d}\|J_T\|_{l^2}^{l}\max_{|\beta|=l}\|\partial^{\beta}w\|_{0,\infty,T}
\end{aligned}$$

den Beweis. Da $F_T : \hat{T} \to T$ bijektiv ist, folgt die zweite Gleichung entsprechend. □

Lemma 4.2.2 *Mit den Bezeichnungen dieses Kapitels gilt*

$$|\det(J_T)| = \frac{|T|}{|\hat{T}|}, \quad \|J_T\|_{l^2} \leq \frac{h_T}{\rho_{\hat{T}}} \text{ und } \|J_T^{-1}\|_{l^2} \leq \frac{h_{\hat{T}}}{\rho_T}. \tag{4.7}$$

Beweis [EG04, Lemma 1.100]. □

Korollar 4.2.3 *Sei $\{\mathcal{T}_h^n\}_{h>0}$ eine lokal quasiuniforme Familie von affinen Partitionierungen mit dem Referenzelement $\hat{T} = (-1,1)^d$. Dann gilt*

$$\|J_T\|_{l^2} \leq \frac{h_T}{2}, \quad \|J_T^{-1}\|_{l^2} \leq \frac{2\sqrt{d}\sigma_0}{h_T} \tag{4.8}$$

und

$$\begin{aligned}
\det(J_T) = \prod_{i=0}^{d_x}\sqrt{\lambda_i} \leq \sqrt{\lambda_{\max}}^{d} = \|J_T\|_{l^2}^{d} \leq 2^{-d}h_T^{d}, \\
|T| = |\hat{T}||\det(J_T)| \leq h_T^{d},
\end{aligned} \tag{4.9}$$

wobei λ_i, $0 \leq \lambda_i \leq d_x$ die Eigenwerte der Matrix $J_T^T J_T$ sind.

Unter diesen Voraussetzungen kann der *Lagrangesche Interpolationsoperator* wie folgt definiert werden:

$$I_h^k : C^0(\overline{T}) \ni v \mapsto I_h^k v = \sum_{i=1}^{n_{\text{dof}}^k} v(x_i)\varphi_i \in \mathbb{Q}_k(T) \subset W^{l,\infty}(T), \quad x_i \in \mathcal{N}.$$

Die Ergebnisse aus diesem Abschnitt über den Lagrangeschen Interpolationsoperator stellen keine Bedingungen an die Lage der Knotenpunkte $x_i \in \mathcal{N}$. Später werden Aussagen über den Interpolationsoperator getroffen, die die speziellen Eigenschaften der Gauss-Lobatto Quadraturpunkte nutzen.

Durch Nachrechnen können die Eigenschaften

$$I_h^k v = v, \quad \forall v \in \mathbb{Q}_k(T), \tag{4.10}$$
$$\|I_h^k v\|_{l,p,T} \leq C\|v\|_{0,\infty,T},\ C > 0,\ l \geq 0,\ p \in [0,\infty] \tag{4.11}$$

bestätigt werden. Ferner gelten die Fehlerabschätzungen:

Lemma 4.2.4 *Es sei T ein Element einer affinen Partitionierung $\mathcal{T}_h^n$, deren zugehörige Familie von Partitionierungen $\{\mathcal{T}_h^n\}_{h>0}$ lokal quasiuniform ist und $1 \leq l \leq k+1$, $l \in \mathbb{N}_0$, $p \in [1,\infty]$ derart, dass $lp > d$ gilt. Dann existieren für den Lagrangeschen Interpolationsoperator I_h^k Konstanten $C > 0$ unabhängig von h mit*

$$|v - I_h^k v|_{r,p,T} \leq C h_T^{l-r} |v|_{l,p,T}, \quad r \leq l, \tag{4.12}$$
$$|v - I_h^k v|_{r,p,E} \leq C h_T^{l-1/p-r} |v|_{l,p,T}, \quad 1/p + r < l, \tag{4.13}$$

für alle $v \in W^{l,p}(T)$, $0 \leq r$. E bezeichne hierbei eine Kante von T.

Beweis Für $1 \leq l \leq k+1$, $r \in \mathbb{N}_0$, $r \leq l$ gilt mit [EG04, Theorem B. 46] für $lp > d$

$$\begin{aligned}\|\hat{v} - I_h^k \hat{v}\|_{r,p,\hat{T}} &\leq \|\hat{v}\|_{r,p,\hat{T}} + \|I_h^k \hat{v}\|_{r,p,\hat{T}} \underset{(4.11)}{\leq} \|\hat{v}\|_{l,p,\hat{T}} + C\|\hat{v}\|_{0,\infty,\hat{T}} \\ &\leq C\|\hat{v}\|_{l,p,\hat{T}}, \quad \forall v \in W^{l,p}(\hat{T}),\end{aligned} \tag{4.14}$$

so dass als weitere Konsequenz folgt

$$\begin{aligned}|(I - I_h^k)\hat{v}|_{r,p,\hat{T}} &\leq \|(I - I_h^k)\hat{v}\|_{r,p,\hat{T}} \underset{(4.10)}{=} \inf_{\hat{p} \in \mathbb{Q}_k(\hat{T})} \|(I - I_h^k)(\hat{v} + \hat{p})\|_{r,p,\hat{T}} \\ &\underset{l-1 \leq k}{\leq} \inf_{\hat{p} \in \mathbb{Q}_{l-1}(\hat{T})} \|(I - I_h^k)(\hat{v} + \hat{p})\|_{r,p,\hat{T}} \\ &\leq \|I - I_h^k\|_{\mathcal{L}(W^{l,p}(\hat{T}), W^{r,p}(\hat{T}))} \inf_{\hat{p} \in \mathbb{Q}_{l-1}(\hat{T})} \|\hat{v} + \hat{p}\|_{l,p,\hat{T}} \\ &\underset{(4.14)}{\leq} C \inf_{\hat{p} \in \mathbb{Q}_{l-1}(\hat{T})} \|\hat{v} + \hat{p}\|_{l,p,\hat{T}} \leq C|\hat{v}|_{l,p,\hat{T}},\end{aligned} \tag{4.15}$$

wobei die letzte Schlussfolgerung auf Lemma 2.3.4 beruht. Die Anwendung des Lemmas 4.2.1 liefert die Transformation auf das physikalische Element T

$$\begin{aligned}|v - I_h^k v|_{r,p,T} &\leq C\|J_T^{-1}\|_{l^2}^r |\det(J_T)|^{1/p} |\hat{v} - I_h^k \hat{v}|_{r,p,\hat{T}} \\ &\leq C\|J_T^{-1}\|_{l^2}^r |\det(J_T)|^{1/p} |\hat{v}|_{l,p,\hat{T}} \\ &\leq C\left(\|J_T\|_{l^2}\|J_T^{-1}\|_{l^2}\right)^r \|J_T\|_{l^2}^{l-r} |v|_{l,p,T} \\ &\leq C\left(\frac{h_T}{\rho_T}\right)^r h_T^{l-r} |v|_{l,p,T} \underset{(4.3)}{\leq} C h_T^{l-r} |v|_{l,p,T}.\end{aligned}$$

Ist $r \in \mathbb{R} \setminus \mathbb{N}_0$, so erzeugt sich mit (2.29)

$$\begin{aligned}|v - I_h^k v|_{r,p,T} &\leq C|v - I_h^k v|_{\lfloor r \rfloor,p,T}^{1-\theta} |v - I_h^k v|_{\lceil r \rceil,p,T}^{\theta},\ \theta = r - \lfloor r \rfloor \\ &\leq C h_T^{l-r} |v|_{l,p,T},\ \lceil r \rceil \leq l.\end{aligned}$$

Die Fehlerabschätzung auf dem Rand erfolgt fast analog. Für $1 \leq l \leq k+1$, $r \in \mathbb{R}_+$ liefert das Spurtheorem 2.2.9 mit der Bemerkung 2.2.10 für jede Kante $\hat{E}$ des Referenzelements $\hat{T}$ mit $E = F_T\hat{E}$ und $0 < \tilde{r} \leq l - 1/p$ somit

$$\begin{aligned}\|\hat{v} - I_h^k\hat{v}\|_{0,p,\hat{E}} \leq C\|\hat{v} - I_h^k\hat{v}\|_{\tilde{r},p,\hat{E}} &\leq C\|\hat{v} - I_h^k\hat{v}\|_{\tilde{r}+1/p,p,\hat{T}} \\ &\underset{(4.14)}{\leq} C\|\hat{v}\|_{l,p,\hat{T}}, \quad \forall \hat{v} \in W^{l,p}(\hat{T}),\end{aligned} \tag{4.16}$$

bzw. mit $-\epsilon < \underbrace{\tilde{r} - \epsilon}_{=r} \leq l - 1/p - \epsilon < l - 1/p$, $0 < \epsilon \leq \tilde{r}$

$$\|\hat{v} - I_h^k\hat{v}\|_{r,p,\hat{E}} \leq C\|\hat{v} - I_h^k\hat{v}\|_{r+\epsilon,p,\hat{E}} \leq C\|\hat{v}\|_{l,p,\hat{T}}, \quad 0 \leq r < l - 1/p.$$

Ferner gilt

$$\begin{aligned}|(I - I_h^k)\hat{v}|_{r,p,\hat{E}} \leq \|(I - I_h^k)\hat{v}\|_{r,p,\hat{E}} &\underset{(4.10)}{=} \inf_{\hat{p}\in\mathbb{Q}_k(\hat{T})} \|(I - I_h^k)(\hat{v} + \hat{p})\|_{r,p,\hat{E}} \\ &\underset{l-1\leq k}{\leq} \inf_{\hat{p}\in\mathbb{Q}_{l-1}(\hat{T})} \|(I - I_h^k)(\hat{v} + \hat{p})\|_{r,p,\hat{E}} \\ &\leq \|I - I_h^k\|_{\mathcal{L}(W^{l,p}(\hat{T}),W^{r,p}(\hat{E}))} \inf_{\hat{p}\in\mathbb{Q}_{l-1}(\hat{T})} \|\hat{v} + \hat{p}\|_{l,p,\hat{T}} \\ &\underset{(4.16)}{\leq} C \inf_{\hat{p}\in\mathbb{Q}_{l-1}(\hat{T})} \|\hat{v} + \hat{p}\|_{l,p,\hat{T}} \leq C|\hat{v}|_{l,p,\hat{T}}.\end{aligned} \tag{4.17}$$

Die Transformation auf T erfolgt unter der Beachtung $h_T|E| \leq C|T|$ ebenfalls analog.

$$\begin{aligned}|v - I_h^k v|_{r,p,E} &\underset{(2.29)}{\leq} C|v - I_h^k v|_{\lfloor r\rfloor,p,E}^{1-\theta}|v - I_h^k v|_{\lceil r\rceil,p,E}^{\theta} \\ &\leq C\|J_E^{-1}\|_{l^2}^r|\det(J_E)|^{1/p}|\hat{v}|_{l,p,\hat{T}} \\ &\leq C\left(\|J_T\|_{l^2}\|J_E^{-1}\|_{l^2}\right)^r \|J_T\|_{l^2}^{l-r}\left(\frac{|E||\hat{T}|}{|\hat{E}||T|}\right)^{1/p}|v|_{l,p,T} \\ &\leq C\left(\frac{h_T}{\rho_E}\right)^r h_T^{l-r}h_T^{-1/p}|v|_{l,p,T} \\ &\underset{(4.3)}{\leq} Ch_T^{l-1/p-r}|v|_{l,p,T}.\end{aligned}$$

□

Bemerkung 4.2.5 Ist $r < 1 - 1/p \leq l - 1/p$ so sichert das Spurtheorem 2.2.9 für $v \in W^{l,p}(T)$ bzw. $v - I_h^k v \in W^{r+1/p,p}(T)$

$$(v - I_h^k v)|_{\partial T} \in W^{r,p}(\partial T).$$

Außerdem gilt dann mit $\|v - I_h^k v\|_{r,p,\partial T}^p = \sum_{E\in\partial T} \|v - I_h^k v\|_{r,p,E}^p$ eine Abschätzung der Form

$$|v - I_h^k v|_{r,p,\partial T} \leq Ch_T^{l-1/p-r}|v|_{l,p,T}.$$

Korollar 4.2.6 *Es sei T ein Element einer affinen Partitionierung $\mathcal{T}_h^n$, deren zugehörige Familie von Partitionierungen $\{\mathcal{T}_h^n\}_{h>0}$ lokal quasiuniform ist und $1 \leq l \in \mathbb{N}_0$, $p \in [1,\infty]$ derart,*

dass $lp > d$ gilt. Dann existieren für den Lagrangeschen Interpolationsoperator I_h^k Konstanten $C > 0$ unabhängig von h mit

$$|v - I_h^k v|_{r,p,T} \leq C h_T^{\min\{k+1,l\}-r} \|v\|_{l,p,T}, \quad r \leq l, \tag{4.18}$$

$$|v - I_h^k v|_{r,p,E} \leq C h_T^{\min\{k+1,l\}-1/p-r} \|v\|_{l,p,T}, \quad 1/p + r < l, \tag{4.19}$$

für alle $v \in W^{l,p}(T)$, $0 \leq r$.

Beweis Für $1 \leq l \leq k+1$ ist nichts zu zeigen. Sei jetzt $k+1 < l$, $r \in \mathbb{R}$. Aus dem Lemma 4.2.4 folgt die Ungleichung

$$|v - I_h^k v|_{r,p,T} \leq C h_T^{k+1-r} |v|_{k+1,p,T}$$

und mit $h_T^{k+1-r}|v|_{k+1,p,T} \leq C h_T^{\min\{k+1,l\}-r}\|v\|_{l,p,T}$ die Behauptung. Der Rest des Beweises verläuft analog. □

Eine weitere Möglichkeit Funktionen in Sobolev-Räumen zu approximieren, stellt die Projektion bzgl. des L^2-Innenproduktes dar:

Definition 4.2.7 Die orthogonale L^2-*Projektion* $P_h^k : L^2(T) \to \mathbb{Q}_k(T)$ wird für $v \in L^2(T)$ definiert durch

$$(v - P_h^k v, w)_{0,T} = 0, \quad \forall w \in \mathbb{Q}_k(T). \tag{4.20}$$

Die Lösung des resultierenden Gleichungssystems zur Bestimmung der L^2-Projektion, lässt sich vermeiden, sobald eine orthogonale Basis existiert. Vermöge der affinen Partitionierung gilt

$$\int_T \psi_\alpha \psi_\beta \, dx = |\det(J_T)| \int_{\hat{T}} \hat{\psi}_\alpha \hat{\psi}_\beta \, d\hat{x} = |\det(J_T)| \prod_{i=0}^{d_x} \frac{1}{2\alpha_i + 1} \delta_{\alpha\beta} = \rho_\alpha^J \delta_{\alpha,\beta}, \tag{4.21}$$

für einen Multiindex α, $\|\alpha\|_{l^\infty} \leq k$ und der Tensorproduktbasis aus *Legendre-Polynomen*

$$\hat{\psi}_\alpha(\hat{x}) = \hat{\psi}_{\alpha_0}^{\alpha_0}(\hat{x}_0)\hat{\psi}_{\alpha_1}^{\alpha_1}(\hat{x}_1)\cdots\hat{\psi}_{\alpha_{d_x}}^{\alpha_{d_x}}(\hat{x}_{d_x}),$$

wobei

$$\hat{\psi}_{\alpha_i}^{\alpha_i}(\hat{x}_i) = \frac{1}{2^{\alpha_i}\alpha_i!} \frac{d^{\alpha_i}}{d\hat{x}_i^{\alpha_i}} (\hat{x}_i + 1)^{\alpha_i}(\hat{x}_i - 1)^{\alpha_i}$$

das α_i-te eindimensionale Legendre-Polynom vom Grad α_i bzgl. $-1 \leq \hat{x}_i \leq 1, 0 \leq i \leq d_x$ ist. In diesem Fall und entsprechender Indizierung ist $P_h^k v = \sum_{i=1}^{n_{\text{dof}}^k} (\rho_i^J)^{-1}(v, \psi_i)_{0,T}\psi_i$.

Zusammenfassende Darstellungen über Legendre-Polynome finden sich in [KS05, Appendix A] und [QV94, Chapter 4].

Aus der Definition der L^2-Projektion ergeben sich die Eigenschaften:

$$P_h^k v = v, \quad \forall v \in \mathbb{Q}_k(T), \tag{4.22}$$

$$\|P_h^k v\|_{0,2,T} \leq \|v\|_{0,2,T}, \quad \forall v \in W^{0,2}(T). \tag{4.23}$$

Lemma 4.2.8 *Es sei T ein Element einer affinen Partitionierung $\mathcal{T}_h^n$, deren zugehörige Familie von Partitionierungen $\{\mathcal{T}_h^n\}_{h>0}$ lokal quasiuniform ist. Dann existieren für $0 \leq l \leq k+1$, $l \in \mathbb{N}_0$ Konstanten $C > 0$ unabhängig von h und k mit*

$$|v - P_h^k v|_{r,2,T} \leq C \frac{h_T^{l-r}}{k^{e(r,l)}} |v|_{l,2,T}, \quad r \leq l, \tag{4.24}$$

$$|v - P_h^k v|_{r,2,E} \leq C \frac{h_T^{l-1/2-r}}{k^{e(r+1/2+\epsilon,l)}} |v|_{l,2,T}, \quad 1/2 + r < l,\ 0 < \epsilon \ll 1, \tag{4.25}$$

und

$$e(r,l) = \begin{cases} l + 1/2 - 2r & r \geq 1 \\ l - 3r/2 & 0 \leq r \leq 1 \end{cases} \tag{4.26}$$

für alle $v \in W^{l,2}(T)$, $0 \leq r$.

Beweis Die Existenz eines $C > 0$ unabhängig von k mit

$$\|\hat{v} - P_h^k \hat{v}\|_{r,2,\hat{T}} \leq C k^{-e(r,l)} \|\hat{v}\|_{l,2,\hat{T}}$$

für $r, l \in \mathbb{R}$, $0 \leq r \leq l$ ist durch [CQ82, Theorem 2.4] gesichert. Mit (4.22) folgt analog zum Interpolationsfehler auf dem Referenzelement für $0 \leq l \leq k+1$

$$|(I - P_h^k)\hat{v}|_{r,2,\hat{T}} \leq C k^{-e(r,l)} |\hat{v}|_{l,2,\hat{T}}.$$

Es gilt für ein durch $\hat{v}$ festgelegtes l mit Lemma 2.3.4: $C = C(d, l, \hat{T})$. Die Konstante ist somit unabhängig vom Polynomgrad k.

Auf den Kanten hingegen, führt die Ungleichungskette

$$\begin{aligned} \|\hat{v} - P_h^k \hat{v}\|_{r,2,\hat{E}} &\leq C \|\hat{v} - P_h^k \hat{v}\|_{r+\epsilon,2,\hat{E}} \leq C \|\hat{v} - P_h^k \hat{v}\|_{r+\epsilon+1/2,2,\hat{T}} \\ &\leq C k^{-e(r+\epsilon+1/2,l)} \|\hat{v}\|_{l,2,\hat{T}}, \quad 0 \leq r < l - 1/2 \end{aligned}$$

zu

$$|(I - P_h^k)\hat{v}|_{r,2,\hat{E}} \leq C k^{-e(r+\epsilon+1/2,l)} |\hat{v}|_{l,2,\hat{T}}.$$

Die abschließende Transformation auf T verläuft analog zum Interpolationsfehler. □

Bemerkung 4.2.9 Speziell ergibt sich für die L^2-Norm (vgl. nachfolgendes Lemma) die Abschätzung:

$$|v - P_h^k v|_{0,2,E} \leq C \frac{h_T^{l-1/2}}{k^{l-3/2(1/2+\epsilon)}} |v|_{l,2,T} \underset{\epsilon<1/6}{\leq} C \frac{h_T^{l-1/2}}{k^{l-1}} |v|_{l,2,T}. \tag{4.27}$$

Bemerkung 4.2.10 Die Bemerkung 4.2.5 und das Korollar 4.2.6 besitzen ihr entsprechendes Analogon.

Im Unterschied zu den eben präsentierten Projektionsfehlerabschätzungen, kann die nachstehende Aussage auch ohne die Verwendung des Lemmas von Deny-Lions gezeigt werden.

Lemma 4.2.11 *Es sei T Element einer affinen Partitionierung $\mathcal{T}_h^n$, deren zugehörige Familie von Partitionierungen $\{\mathcal{T}_h^n\}_{h>0}$ lokal quasiuniform ist. Dann existieren Konstanten unabhängig von h und k, derart, dass mit $0 \leq l$ und $v \in W^{l,2}(T)$ für $1 \leq s \leq \min\{k+1, l\}$ gilt:*

$$|v - P_h^k v|_{0,2,T} \leq C \left(\frac{h_T}{k}\right)^s |v|_{s,2,T}, \tag{4.28}$$

$$|v - P_h^k v|_{1,2,T} \leq C \frac{h_T^{s-1}}{k^{s-3/2}} |v|_{s,2,T}, \tag{4.29}$$

$$|v - P_h^k v|_{0,2,\partial T} \leq C \frac{h_T^{s-1/2}}{k^{s-1}} |v|_{s,2,T}. \tag{4.30}$$

Beweis [Geo03, Corollary 3.15, 3.19, (3.1)-(3.3)], [SVD06, Lemma 6.1, Remark 6.2]. □

Bemerkung 4.2.12 Die Abschätzungen (4.27) und (4.30) sind beide suboptimal von der Größenordnung $k^{1/2}$. Der Grund ist das nicht optimale Verhalten des L^2-Projektionsfehlers in $|\cdot|_{r,2,T}$ für $r > 0$. Beruht der Beweis der genannten Abschätzungen auf einer, wie auch immer gearteten Spuraussage, so entstehen zwangsläufig auf der rechten Seite Normen in denen das nichtoptimale Verhalten eine Rolle spielt.

4.3 Quadratur und Lumping

Zur näherungsweisen Berechnung der auftretenden Integrale werden interpolatorische Quadraturformeln betrachtet (vgl. [EG04, Definition 8.1]):

Definition 4.3.1 Sei $T \subset \mathbb{R}^{d_x}$ ein nichtleeres, zusammenhängendes und kompaktes Lipschitz-Gebiet. Eine Quadraturformel mit n_{dof}^k Punkten besteht aus:

1. Einer Menge aus n_{dof}^k reellen Zahlen $\left\{\omega_1^J, \ldots, \omega_{n_{\text{dof}}^k}^J\right\}$, genannt *Integrationsgewichte*.
2. Einer Menge $\mathcal{Q}$ aus n_{dof}^k Punkten $\{x_1, \ldots, x_{n_{\text{dof}}^k}\}$ in T mit $x_i \neq x_j$ falls $i \neq j$, genannt *Quadraturpunkte*.

Die größte natürliche Zahl k, für die

$$\int_T p(x)\,dx = \sum_{i=1}^{n_{\text{dof}}^k} \omega_i^J p(x_i) \quad \forall p \in \mathbb{Q}_k(T) \tag{4.31}$$

gilt, heißt *Genauigkeitsgrad* der Quadraturformel.

Aus der zugrundeliegenden Idee

$$\begin{aligned}
\int_T p(x)\,dx &= \int_{\hat{T}} p(F_T(\hat{x}))|\det(J_T)|\,d\hat{x} \\
&= \int_{\hat{T}} \sum_{\alpha \in \mathbb{N}_0^d,\, \|\alpha\|_{l^\infty} \leq k} p(F_T(\hat{x}_\alpha))\hat{\varphi}_\alpha(\hat{x})|\det(J_T)|\,d\hat{x} \\
&= \sum_{\alpha \in \mathbb{N}_0^d,\, \|\alpha\|_{l^\infty} \leq k} \int_{\hat{T}} \hat{\varphi}_\alpha(\hat{x})|\det(J_T)|\,d\hat{x}\, p(x_\alpha) = \sum_{\alpha \in \mathbb{N}_0^d,\, \|\alpha\|_{l^\infty} \leq k} \omega_\alpha^J p(x_\alpha), \quad x_\alpha \in \mathcal{Q},
\end{aligned}$$

ergeben sich die Integrationsgewichte

$$\omega_i^J = \int_{\hat{T}} \hat{\varphi}_i(\hat{x}) |\det(J_T)| \, d\hat{x} = \int_T \varphi_i(x) \, dx, \quad 1 \leq i \leq n_{\text{dof}}^k,$$

zu den Quadraturpunkten $x_i = F_T(\hat{x}_i)$, wobei $\hat{\varphi}_i$ die Lagrange-Basis zu den Punkten $\hat{x}_i$ ist. Um eine gute Konditionierung des Quadraturverfahrens zu gewährleisten, wird

$$\omega_i^J > 0 \text{ für } 1 \leq i \leq n_{\text{dof}}^k \tag{4.32}$$

gefordert.

Die Definition 4.3.1 der nodalen Quadraturformel kann zur Definition eines diskreten Innenproduktes genutzt werden. Mit der daraus resultierenden Norm können für diskrete Argumente Normäquivalenzabschätzungen bzgl. der L^p-Norm angegeben werden, deren Konstanten nur von p, dem Polynomgrad k und von der Verteilung der Quadraturpunkte abhängen. Um solche Ungleichungen angeben zu können, werden *Kontrollvolumen* konstruiert, die die Definition eines *Lumpingoperators* erlauben. Wie angedeutet wird die diskrete L^2-Norm definiert als

$$\begin{aligned} \|v\|_{0,2,T,h}^2 = (v,v)_{0,T,h} &= \sum_{\alpha \in \mathbb{N}_0^d, \, \|\alpha\|_{l^\infty} \leq k} \omega_\alpha^J \, v(x_\alpha)^2 \\ &= \sum_{\alpha_0=0}^{k} \cdots \sum_{\alpha_{d_x}=0}^{k} \omega_{\alpha_0}^J \cdots \omega_{d_x}^J v(x_\alpha)^2, \quad x_\alpha \in \mathcal{Q}, \end{aligned} \tag{4.33}$$

mit $\sum_{\alpha \in \mathbb{N}_0^d, \, \|\alpha\|_{l^\infty} \leq k} \omega_\alpha^J = |T|$, $\omega_\alpha > 0$, $\forall \alpha \in \mathbb{N}_0^d$. Die Verallgemeinerung für die L^p-Norm geschieht analog.

Die Konstruktion der Kontrollvolumen erfolgt durch

$$\Omega_\alpha = \left(\sum_{i=0}^{\alpha_0 - 1} \omega_i^J, \sum_{i=0}^{\alpha_0} \omega_i^J \right) \times \cdots \times \left(\sum_{i=0}^{\alpha_{d_x} - 1} \omega_i^J, \sum_{i=0}^{\alpha_{d_x}} \omega_i^J \right), \tag{4.34}$$

so dass sich für das Kontrollvolumen

$$|\Omega_\alpha| = \omega_{\alpha_0}^J \omega_{\alpha_1}^J \cdots \omega_{\alpha_{d_x}}^J = \omega_\alpha^J \tag{4.35}$$

und

$$\sum_{\alpha \in \mathbb{N}_0^d, \, \|\alpha\|_{l^\infty} \leq k} |\Omega_\alpha| = |T| \tag{4.36}$$

ergibt. Der Lumpingoperator wird nun definiert als $L_h = C(\overline{T}) \to L^\infty(T)$ mit

$$L_h v = \sum_{i=1}^{n_{\text{dof}}^k} v(x_i) \chi_{\Omega_i}, \quad x_i \in \mathcal{Q}, \tag{4.37}$$

wobei

$$\chi_G(x) = \begin{cases} 0 & x \notin G, \\ 1 & x \in G. \end{cases}$$

Die Integration der p-ten Potenz des Lumpingoperators bringt mit (4.35), (4.36) und der affinen Transformation des Referenzelements folgende Identität zum Vorschein

$$\begin{aligned}\|L_h(v)\|_{0,p,T}^p &= |T| \int_{\hat{T}} |L_h(v)|^p \, d\hat{x} = \sum_{j=1}^{n_{\text{dof}}^k} |\Omega_j| \int_{\hat{T}} |L_h(v)|^p \, d\hat{x} \\ &= \sum_{j=1}^{n_{\text{dof}}^k} \int_{\Omega_j} \left| \sum_{i=0}^{n_{\text{dof}}^k} v(x_i) \chi_{\Omega_i}(x) \right|^p dx = \sum_{j=1}^{n_{\text{dof}}^k} |\Omega_j| |v(x_j)|^p \\ &= \|v\|_{0,p,T,h}^p, \quad x_i \in \mathcal{Q}.\end{aligned} \tag{4.38}$$

Ferner ist

$$\|v\|_{0,p,T,h}^p \underset{\mathcal{N}=\mathcal{Q}}{=} \int_T \sum_{j=1}^{n_{\text{dof}}^k} |v(x_j)|^p \varphi_j(x) \, dx = \int_T I_h^k(|v|^p) \, dx. \tag{4.39}$$

Lemma 4.3.2 *Es existieren Konstanten $C_{L_1}, C_{L_2} > 0$ unabhängig von h_T, so dass die Äquivalenzabschätzungen*

$$C_{L_1} \|v\|_{0,p,T} \leq \|L_h(v)\|_{0,p,T} \leq C_{L_2} \|v\|_{0,p,T}, \quad \forall v \in \mathbb{Q}_k(T) \tag{4.40}$$

gelten.

Beweis Die Unabhängigkeit der Konstanten von h_T ergibt sich für alle T, die durch eine affine Transformation aus dem Referenzelement hervorgegangen sind, direkt aus der Substitutionsregel. Aus $\|L_h(v)\|_{0,p,T}^p = \sum_{j=1}^{n_{\text{dof}}^k} |\Omega_j| |v(x_j)|^p$ folgt sofort

$$\begin{aligned}\|v\|_{0,p,\hat{T}} &\underset{(4.62)}{\leq} (3k^2)^{d\frac{p-2}{2p}} \|v\|_{0,2,\hat{T}} \leq (3k^2)^{d\frac{p-2}{2p}} \lambda_{\max}(\hat{M})^{1/2} \|v_{\mathcal{N}}\|_{l^2} \\ &\underset{(2.6)}{\leq} (3k^2(k+1))^{d\frac{p-2}{2p}} \lambda_{\max}(\hat{M})^{1/2} \|v_{\mathcal{N}}\|_{l^p} \\ &\leq (3k^2(k+1))^{d\frac{p-2}{2p}} \lambda_{\max}(\hat{M})^{1/2} \left(\min_{1 \leq j \leq n_{\text{dof}}^k} |\Omega_j| \right)^{-1/p} \|L_h(v)\|_{0,p,\hat{T}}\end{aligned}$$

und

$$\begin{aligned}\|L_h(v)\|_{0,p,\hat{T}} &\leq \left(\max_{1 \leq j \leq n_{\text{dof}}^k} |\Omega_j| \right)^{1/p} \|v\|_{l^p} \leq \left(\max_{1 \leq j \leq n_{\text{dof}}^k} |\Omega_j| \right)^{1/p} \|v\|_{l^2} \\ &\leq \left(\max_{1 \leq j \leq n_{\text{dof}}^k} |\Omega_j| \right)^{1/p} \|v\|_{0,2,\hat{T}} \lambda_{\min}(\hat{M})^{-1/2} \\ &\leq |\hat{T}|^{(p-2)/2p} \left(\max_{1 \leq j \leq n_{\text{dof}}^k} |\Omega_j| \right)^{1/p} \|v\|_{0,p,\hat{T}} \lambda_{\min}(\hat{M})^{-1/2},\end{aligned}$$

wobei $\Lambda_{\min}(\hat{M})$ und $\Lambda_{\max}(\hat{M})$ die entsprechenden Eigenwerte der zugehörigen *Massenmatrix* $\hat{M}_{ij} = \int_{\hat{T}} \hat{\varphi}_i \hat{\varphi}_j \, d\hat{x}$ sind. □

Für die Approximation des L^2-Projektors, d.h. für die Berechnung der auftretenden Integrale mit Quadraturformeln, ist die Frage nach der Wahl der Quadraturpunkte von entscheidender Bedeutung.

Ferner kann auch bei Verwendung der Lagrangeschen Basispolynome $\hat{\varphi}_i^k$ für $0 \leq i \leq k$ Einfluss auf die Punktwahl genommen werden. Sind beide Punktmengen $\mathcal{N}$ und $\mathcal{Q}$ identisch, so folgt aus der Definition des L^2-Projektors und der Lagrangeschen Basispolynome

$$0 \stackrel{!}{=} (v - P_h^k v, w)_{0,T,h} = \sum_{i=1}^{n_{\text{dof}}^k} \omega_i^J \left\{ v(x_i) - (P_h^k v)(x_i) \right\} w(x_i), \quad \forall w \in \mathbb{Q}_k(T) \tag{4.41}$$

die Bedingung $v(x_i) = (P_h^k v)(x_i)$ für alle Quadraturpunkte. Mit der Darstellung des Projektors $(P_h^k v)(x) = \sum_{j=1}^{n_{\text{dof}}^k} (P_h^k v)_j \varphi_j(x)$ ergibt sich $(P_h^k v)_i = v(x_i)$. Bei der Approximation der L^2-Projektion gilt somit $\mathcal{N} = \mathcal{Q} \Rightarrow P_h^k = I_h^k$.

Die höchste Genauigkeit liefert die *Gauss-Quadratur*. Sie ist bei $(k+1)^d$ Quadraturpunkten exakt für Polynome vom Grad $2k+1$ (vgl. z.B. [EG04, Proposition 8.2], [BM97, S. 294 ff.] oder [Can+07, S. 448 ff.]). Die Quadraturpunkte in $(-1, 1)$ ergeben sich aus den Nullstellen des $(k+1)$-ten Legendre-Polynoms $\hat{\psi}_{k+1}^{k+1}(x)$, $-1 \leq x \leq 1$. Allerdings gelten für den Interpolationsfehler bzgl. der Gauss-Quadraturpunkte nicht immer optimale Abschätzungen ([BM97, (13.15)]):

Lemma 4.3.3 *Für alle reellen Zahlen r und l, $r \leq l$, $l \geq 1$ existiert ein $C > 0$ unabhängig von k mit*

$$\|\hat{v} - I_G^k \hat{v}\|_{r,2,I} \leq C k^{-e(r,l)} \|\hat{v}\|_{l,2,I}, \quad \forall \hat{v} \in W^{l,2}(I)$$

und $e(r,l)$ aus dem Lemma 4.2.8.

Eine Alternative besteht in der Wahl der Quadraturpunkte nach *Gauss-Lobatto*. Sie sind als die Nullstellen des Polynoms

$$(x+1)(x-1)(\hat{\psi}_k^k)'(x), \; -1 \leq x \leq 1 \tag{4.42}$$

definiert und ergeben eine exakte Quadratur für Polynome vom Grad $2k-1$ (vgl. [Can+07, S. 448]). Offensichtlich sind die Randpunkte stets Quadraturpunkte. Im Unterschied zum vorherigen Lemma gilt für die *Gauss-Lobatto-Interpolation*:

Lemma 4.3.4 *Für die reellen Zahlen r und l mit $2l > d + r$ und $0 \leq r \leq 1$ existiert ein C unabhängig von k, so dass die folgende Abschätzung gilt*

$$\|\hat{v} - I_{GL}^k \hat{v}\|_{r,2,\hat{T}} \leq C k^{r-l} \|\hat{v}\|_{l,2,\hat{T}}, \quad \forall \hat{v} \in W^{l,2}(\hat{T}). \tag{4.43}$$

Beweis [BM97, Theorem 14.2]. □

Bei der entsprechenden Quadratur gilt das folgende Resultat (vgl. Lemma 4.3.2).

Lemma 4.3.5 *Sei $\mathcal{Q}$ die Menge der Gauss-Lobatto-Quadraturpunkte. Dann existiert eine Konstante C unabhängig von h_T und k, so dass die Äquivalenzabschätzungen*

$$\|v\|_{0,2,T} \leq \|L_h(v)\|_{0,2,T} \leq C\|v\|_{0,2,T},\ \forall v \in \mathbb{Q}_k(T) \tag{4.44}$$

gelten.

Beweis Die affine Transformation auf das Referenzelement liefert die Unabhängigkeit der Konstante von h_T. Der Beweis für das Referenzelement erfolgt mit $\|L_h(v)\|^2_{0,2,\hat{T}} \underset{(4.38)}{=} \|v\|^2_{0,2,\hat{T},GL}$ nach [CQ82, (3.9)]. □

Bemerkung 4.3.6 Analog zu Lemma 4.3.2 kann eine Ungleichungskette für $p \neq 2$ gefolgert werden.

Eine weitere interessante Eigenschaft der Gauss-Lobatto-Quadraturpunkte kommt in der Quadratsumme der Lagrange-Polynome zum Vorschein:

Lemma 4.3.7 *Seien $\hat{\varphi}_\alpha(\hat{x})$ die Lagrange-Polynome bzgl. der Gauss-Lobatto-Quadraturpunkte. Dann gilt*

$$\sum_{\substack{\alpha \in \mathbb{N}_0^d, \\ \|\alpha\|_{l^\infty} \leq k}} \hat{\varphi}_\alpha(\hat{x})^2 \leq 1, \quad \forall \hat{x} \in \hat{T}. \tag{4.45}$$

Beweis Aus der Tensorproduktdarstellung folgt mit [Fej32, §1] direkt die Behauptung:

$$\sum_{\substack{\alpha \in \mathbb{N}_0^d, \\ \|\alpha\|_{l^\infty} \leq k}} \hat{\varphi}_\alpha(\hat{x})^2 = \underbrace{\sum_{\alpha_0=0}^{k} \hat{\varphi}^k_{\alpha_0}(\hat{x}_0)^2}_{\leq 1} \cdots \underbrace{\sum_{\alpha_{d_x}=0}^{k} \hat{\varphi}^k_{\alpha_{d_x}}(\hat{x}_{d_x})^2}_{\leq 1} \leq 1. \qquad \square$$

4.4 Projektions- und Interpolationsfehler bzgl. der Gauss-Lobatto-Quadraturpunkte

Die zur Darstellung des L^2-Projektors verwendeten orthogonalen Legendre-Polynome ψ_i für $1 \leq i \leq n^k_{\text{dof}}$ besitzen zwei nützliche Eigenschaften: Zum einen bilden sie eine orthogonale Basis von $\mathbb{Q}_k(T)$, so dass die Massenmatrix Diagonalgestalt annimmt. Des Weiteren sorgen die Polynome für eine hierarchische Zerlegung von $\mathbb{Q}_k(T)$ im Sinne von:

Definition 4.4.1 (Hierarchische modale Basis) Eine Familie $\{\mathcal{B}_k\}_{k\geq 0}$, wobei $\mathcal{B}_k$ eine Menge von Polynomen ist, wird genau dann hierarchische modale Basis genannt, wenn für alle $k \geq 0$ die folgenden Eigenschaften erfüllt sind:

1. $\mathcal{B}_k$ ist eine Basis von $\mathbb{Q}_k$,
2. $\mathcal{B}_k \subset \mathcal{B}_{k+1}$.

Die hierarchische Basis, bestehend aus Legendre-Polynomen ψ_i, $1 \leq i \leq n_{\text{dof}}^k$, liefert unter Benutzung der Orthogonalitäteigenschaft für die L^2-Projektion von $v \in \mathbb{Q}_k(T)$ z.B.:

$$\left((I - P_h^K)\nabla v, (I - P_h^K)\nabla v\right)_{0,T} = \sum_{i=n_{\text{dof}}^K+1}^{n_{\text{dof}}^k} \left(\frac{(\nabla v, \psi_i)_{0,T}}{\rho_i^J}\right)^2 (\psi_i, \psi_i)_{0,T} \geq 0.$$

Diese Nichtnegativität kann allerdings nicht auf die Argumente v, v^{p-1}, $p = 2m, m \in \mathbb{N}$ verallgemeinert werden:

$$\left((I - P_h^K)\nabla v, (I - P_h^K)\nabla v^{p-1}\right)_{0,T} = \sum_{i=n_{\text{dof}}^K+1}^{n_{\text{dof}}^k} \frac{(\nabla v, \psi_i)_{0,T}}{\rho_i^J} \frac{(\nabla v^{p-1}, \psi_i)_{0,T}}{\rho_i^J} (\psi_i, \psi_i)_{0,T} \not\geq 0.$$

Eine solche Abschätzung ist allerdings für die $L^\infty(L^\infty)$-Analysis der lokalen Projektion des Shock-capturing Terms notwendig, so dass es erforderlich wird nach einem anderen Projektor Ausschau zu halten. Der Rest dieses Abschnittes widmet sich dieser Aufgabe und stellt einige Eigenschaften des neuen Projektors bereit. Die geforderte Bedingung wird dann im Lemma 4.4.6 bewiesen.

Wird die L^2-Projektion mit Lagrangeschen Polynomen und $\mathcal{N} = \mathcal{Q}$ diskretisiert, so erhält sich dank der Hilfe von $\varphi_i(x_j) = \delta_{ij}$, $1 \leq i, j \leq n_{\text{dof}}^k$ die Orthogonalitätseigenschaft der Basis. Aus diesem Grund ist die dazugehörige Massenmatrix ebenfalls eine Diagonalmatrix. Die Ausnutzung von $\sum_{i=1}^{n_{\text{dof}}^k} \varphi_i(x) = 1$, $\forall x \in T$ und ebenfalls $\mathcal{N} = \mathcal{Q}$ macht deutlich, dass diese Diskretisierung genau der Summierung der Zeilen- bzw. Spalteneinträge der exakten Massenmatrix entspricht. Dieser Prozess wird in der Literatur auch als mass lumping bezeichnet und entsteht nahtlos aus

$$(L_h(\varphi_i), L_h(\varphi_j))_{0,T} \underset{(4.38)}{=} (\varphi_i, \varphi_j)_{0,T,h} \underset{\mathcal{N}=\mathcal{Q}}{=} \omega_i^J \delta_{ij}, \quad 1 \leq i, j, \leq n_{\text{dof}}^k. \tag{4.46}$$

Der bei der Diskretisierung erfolgte Übergang der Basis zwischen den Legendre-Polynomen zu den Lagrange-Polynomen rettete zwar speziell für $\mathcal{N} = \mathcal{Q}$ die Orthogonalitätseigenschaft, erfüllt mit den Lagrange-Polynomen aber nicht mehr die Bedingungen der hierarchischen Basis. Allerdings kann folgende Definition erfüllt werden:

Definition 4.4.2 (Eingebettete hierarchische nodale Basis vom Grad K) Eine Familie $\{\mathcal{B}_j\}_{0 \leq j \leq k}$, wobei $\mathcal{B}_j$ eine Menge von Polynomen vom Grad k ist, wird genau dann *eingebettete hierarchische nodale Basis vom Grad* K, $K \geq 0$ genannt, wenn eine Menge von Polynomen $\tilde{\mathcal{B}}_K$ vom Grad K mit den folgenden Eigenschaften existiert:

1. $\tilde{\mathcal{B}}_K$ mit $N(\tilde{\mathcal{B}}_K) \subseteq N(\mathcal{B}_k)$ ist eine Basis von $\mathbb{Q}_K$,
2. $\mathcal{B}_K \subseteq \mathcal{B}_k$, $N(\mathcal{B}_K) = N(\tilde{\mathcal{B}}_K)$,
3. $\mathcal{B}_k$ ist eine Basis von $\mathbb{Q}_k$,

wobei $N : \mathcal{B}_j \ni \varphi \mapsto x \in \mathbb{R}^d$, $0 \leq j \leq k$ die bijektive Funktion ist, die den nodalen Basispolynomen ihre zugeordneten Punkte zuweist.

Beispiel 4.4.3

1. Die Lagrange-Polynome bzgl. der Gauss-Lobatto Quadraturpunkte bilden eine eingebettete hierarchische nodale Basis vom Grad 1 und 2, denn für $d = 1$ gilt $-1, 1 \in N(\mathcal{B}_k)$, $k \geq 1$ und $-1, 0, 1 \in N(\mathcal{B}_k)$, $k = 2, 4, 6, \ldots$, d.h. $N(\tilde{\mathcal{B}}_1) \subseteq N(\mathcal{B}_k)$, $k \geq 1$ und $N(\tilde{\mathcal{B}}_2) \subseteq N(\mathcal{B}_k)$, $k = 2, 4, 6, \ldots$.
2. Die Lagrange-Polynome bzgl. der *Gauss-Kronrod* Quadraturpunkte bilden eine eingebettete hierarchische nodale Basis vom Grad K. Bei den Gauss-Kronrod Quadraturpunkten für $d = 1$ handelt es sich um $K + 1$-Gausspunkte, die um $K + 2$ Punkte ergänzt werden (vgl. z.B. [Cal+00]), so dass gilt

$$\{\tilde{x}_i\}_{i=0}^{K} \subset \{x_i\}_{i=0}^{2K+2}, \tag{4.47}$$

$$\int_0^1 u\,dx = \sum_{i=0}^{2K+2} u(x_i)\omega_i, \quad \forall u \in \mathbb{Q}_{3K+4}(-1, 1). \tag{4.48}$$

   Nach Definition gilt somit $N(\tilde{\mathcal{B}}_K) \subset N(\mathcal{B}_k)$, $K \geq 0$, $k = 2K + 2$.

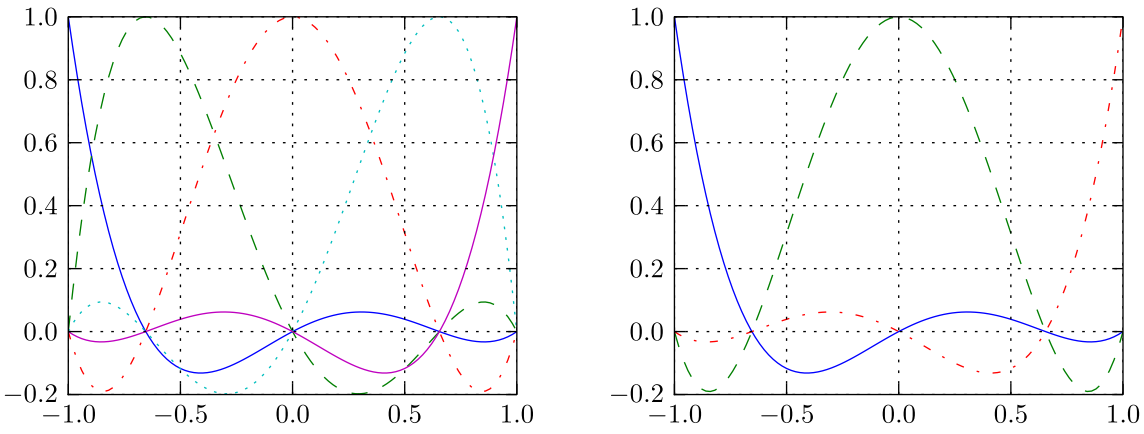
Definition 4.4.4 Sei $\{\mathcal{B}_j\}_{0 \leq j \leq k}$ eine eingebettete hierarchische nodale Basis vom Grad K. Dann ist für $\mathbb{Q}_k(T) = V_h^{K,k}(T) \oplus V_h'(T)$ mit $V_h^{K,k}(T) = \text{span}\{\mathcal{B}_K\}$ und $V_h'(T) = \text{span}\{\mathcal{B}_k \setminus \mathcal{B}_K\}$ der Projektor $P_h^{K,k} : L^2(T) \to V_h^{K,k}(T)$ definiert als

$$(v - P_h^{K,k} v, w)_{0,T,h} = 0, \quad \forall w \in V_h^{K,k}(T). \tag{4.49}$$

In diesem Zusammenhang soll unter dem Fluktuationsoperator der Projektor

$$P_h' = I - P_h^{K,k} \tag{4.50}$$

verstanden werden.



(a) Lagrangesche Basispolynome von $\mathbb{Q}_4(-1, 1)$ zu den Gauss-Lobatto Punkten.

(b) Entsprechende Basispolynome von $V_{GL}^{2,4}(-1, 1)$.

Abbildung 4.2

Es folgen einige Eigenschaften der beiden Projektoren $P_h^{K,k}$ und P_h' :

Lemma 4.4.5 *Sei $\{\mathcal{B}_j\}_{0\le j\le k}$ eine eingebettete hierarchische nodale Basis vom Grad K aus Lagrange-Polynomen und I_h^K der Lagrangesche Interpolationsoperator bzgl. $N(\mathcal{B}_K)$. Dann gilt:*

$$|v - P_h^{K,k}v|_{r,2,T} \le Ch_T^{l-r}\left(|v|_{l,2,T} + |P_h^{K,k}v|_{l,2,T}\right), \quad r \le l,$$
$$|v - P_h^{K,k}v|_{r,2,E} \le Ch_T^{l-1/2-r}\left(|v|_{l,2,T} + |P_h^{K,k}v|_{l,2,T}\right),\ 1/2 + r < l$$

für alle $v \in W^{l,p}(T)$.

Beweis Unter den gegebenen Voraussetzungen gilt für alle Lagrangeschen Basispolynome

$$\varphi_j^{K,k} \in \mathcal{B}_K : \varphi_j^{K,k}(x_i) = \delta_{ij},\ 1 \le i,j \le n_{\text{dof}}^K,\ x_i \in N(\tilde{\mathcal{B}}_K) = N(\mathcal{B}_K).$$

Aus der Definition des Projektors $P_h^{K,k}$ folgt mit

$$w \in V_h^{K,k}(T) \Rightarrow w(x_i) = 0, \quad x_i \in N(\mathcal{B}_k) \setminus N(\mathcal{B}_K) \tag{4.51}$$

die Bedingung

$$0 \overset{!}{=} (v - P_h^{K,k}v, w)_{0,T,h} = \sum_{i=1}^{n_{\text{dof}}^k} \omega_i\{v(x_i) - (P_h^{K,k}v)(x_i)\}w(x_i)$$
$$\underset{(4.51)}{=} \sum_{i=1}^{n_{\text{dof}}^K} \omega_i\{v(x_i) - (P_h^{K,k}v)(x_i)\}w(x_i)$$

bzw.

$$v(x_i) \overset{!}{=} \sum_{j=1}^{n_{\text{dof}}^K} (P_h^{K,k}v)_j \varphi_j^{K,k}(x_i) \underset{\mathcal{N}=\mathcal{Q}}{=} (P_h^{K,k}v)_i,\ 1 \le i \le n_{\text{dof}}^K. \tag{4.52}$$

Als Konsequenz entsteht die Identität

$$I_h^K(P_h^{K,k}v) = \sum_{i=1}^{n_{\text{dof}}^K} (P_h^{K,k}v)(x_i)\varphi_i^K = \sum_{i=1}^{n_{\text{dof}}^K}\sum_{j=1}^{n_{\text{dof}}^K} (P_h^{K,k}v)_j \varphi_j^{K,k}(x_i)\varphi_i^K$$
$$= \sum_{i=1}^{n_{\text{dof}}^K} (P_h^{K,k}v)_i \varphi_i^K \underset{(4.52)}{=} I_h^K v,$$

die bei der Abschätzung des Projektionsfehler im Zusammenhang mit Lemma 4.2.4 gute Dienste leistet:

$$\begin{aligned}\|v - P_h^{K,k}v\|_{r,p,T} &\le \|v - I_h^K v\|_{r,p,T} + \|I_h^K v - P_h^{K,k}v\|_{r,p,T} \\ &\le \|v - I_h^K v\|_{r,p,T} + \|P_h^{K,k}v - I_h^K(P_h^{K,k}v)\|_{r,p,T}.\end{aligned} \tag{4.53}$$

□

Lemma 4.4.6 *Sei $\{\mathcal{B}_j\}_{0\le j\le k}$ eine eingebettete hierarchische nodale Basis vom Grad K aus Lagrange-Polynomen. Dann ist*

1. *$P_h^{K,k}$ linear,*
2. *$\|P_h^{K,k}v\| \le C(K,k,\|\cdot\|)\|v\|_{0,\infty,T}$, $\|\cdot\|$ beliebig,*
3. *$p = 2m, m \in \mathbb{N}$, $v \in \mathbb{Q}_k(T)$:*

$$\hat{\epsilon}_{\mathrm{vms}}(v, v^{p-1}) = \frac{(P_h'(\nabla v), P_h'(\nabla v^{p-1}))_{0,T,h}}{(\nabla v, \nabla v^{p-1})_{0,T}} \ge 0. \tag{4.54}$$

Beweis Zu (i) :

$$P_h^{K,k}(\alpha v + \beta w) \underset{(4.52)}{=} \sum_{i=1}^{n_{\mathrm{dof}}^K} (\alpha v(x_i) + \beta w(x_i))\varphi^{K,k} = \alpha P_h^{K,k} v + \beta P_h^{K,k} w.$$

Zu (ii) :

$$\begin{aligned}\|P_h^{K,k}v\| &\le \sum_{i=1}^{n_{\mathrm{dof}}^K} \|v(x_i)\varphi_i^{K,k}\| \le \sum_{i=1}^{n_{\mathrm{dof}}^K} |v(x_i)|\|\varphi_i^{K,k}\| \\ &\le \left(\sum_{i=1}^{n_{\mathrm{dof}}^K} \|\varphi_i^{K,k}\|\right) \|v\|_{0,\infty,T} \le C(K,k,\|\cdot\|)\|v\|_{0,\infty,T}.\end{aligned}$$

Zu (iii) : Es $\nabla v \cdot \nabla v^{p-1} = \frac{4(p-1)}{p^2}\nabla v^{p/2} \cdot \nabla v^{p/2} \ge 0$, so dass

$$\begin{aligned}(P_h'(\nabla v), P_h'(\nabla v^{p-1}))_{0,T,h} &\underset{(4.49)}{=} (\nabla v, \nabla v^{p-1})_{0,T,h} - (P_h^{K,k}(\nabla v), P_h^{K,k}(\nabla v^{p-1}))_{0,T,h} \\ &\underset{(4.52)}{=} \sum_{i=1}^{n_{\mathrm{dof}}^k} \omega_i^J \nabla v(x_i)\nabla v^{p-1}(x_i) - \sum_{i=1}^{n_{\mathrm{dof}}^K} \omega_i^J \nabla v(x_i)\nabla v^{p-1}(x_i) \\ &= \sum_{i=n_{\mathrm{dof}}^K+1}^{n_{\mathrm{dof}}^k} \omega_i^J \nabla v(x_i)\nabla v^{p-1}(x_i) \\ &= \frac{4(p-1)}{p^2} \sum_{i=n_{\mathrm{dof}}^K+1}^{n_{\mathrm{dof}}^k} \omega_i^J \nabla v^{p/2}(x_i)\nabla v^{p/2}(x_i) \ge 0.\end{aligned}$$

Die Vertauschung von P_h' mit $P_h^{K,k}$ liefert analog $(P_h^{K,k}(\nabla v), P_h^{K,k}(\nabla v^{p-1}))_{0,T,h} \ge 0$. □

Motiviert durch Lemma 4.4.5 und die Fehlerabschätzung (4.43) entsteht die Frage nach dem lokalen Interpolationsfehler des Gauss-Lobatto-Interpolationsoperators.

Lemma 4.4.7 *Es sei T ein Element einer affinen Partitionierung $\mathcal{T}_h^n$, deren zugehörige Familie von Partitionierungen $\{\mathcal{T}_h^n\}_{h>0}$ lokal quasiuniform ist und $1 \leq l \leq k+1$, $l \in \mathbb{N}_0$ derart, dass $2l > d + r$ gilt. Dann existieren für den Lagrangeschen Interpolationsoperator I_{GL}^k Konstanten $C > 0$ unabhängig von h und k mit*

$$|v - I_{GL}^k v|_{r,2,T} \leq C \left(\frac{h_T}{k}\right)^{l-r} |v|_{l,2,T},\ 0 \leq r \leq 1, \tag{4.55}$$

$$|v - I_{GL}^k v|_{r,2,E} \leq C k^{\epsilon} \left(\frac{h_T}{k}\right)^{l-1/2-r} |v|_{l,2,T},\ 0 \leq r,\ 1/2 + r < 1,\ 0 < \epsilon \ll 1 \tag{4.56}$$

für alle $v \in W^{l,p}(T)$.

Beweis Unter Rückgriff auf Lemma 4.3.4 verläuft der Beweis analog zu den Lemmata 4.2.4 und 4.2.8. □

Korollar 4.4.8 *Sei $1 \leq l \leq K+1$. Unter den Voraussetzungen der Lemmata 4.4.5 und 4.4.7 existieren Konstanten $C > 0$ unabhängig von h und k mit*

$$|v - P_{GL}^{K,k} v|_{r,2,T} \leq C \left(\frac{h_T}{K}\right)^{l-r} \left(|v|_{l,2,T} + |P_{GL}^{K,k} v|_{l,2,T}\right),\ 0 \leq r \leq 1, \tag{4.57}$$

$$|v - P_{GL}^{K,k} v|_{r,2,E} \leq C k^{\epsilon} \left(\frac{h_T}{K}\right)^{l-1/2-r} \left(|v|_{l,2,T} + |P_{GL}^{K,k} v|_{l,2,T}\right),\ 0 \leq r, \tag{4.58}$$

$1/2 + r < 1, 0 < \epsilon \ll 1$ *für alle $v \in W^{l,p}(T)$.*

Bemerkung 4.4.9 Die Konvergenzeigenschaft des Projektors $P_{GL}^{K,k}$ wird genau wie im Lemma 4.4.5 auf das Verhalten von $P_{GL}^{K,k}$ in einer Halbnorm zurückgeführt.

Über den Fehler, der durch die Nichtkommutativität von ∇ und I_{GL}^k entsteht, kann folgende Aussage getroffen werden:

Lemma 4.4.10 *Unter den Voraussetzungen von Lemma 4.4.7 existiert ein $C > 0$ unabhängig von h und k mit*

$$\|I_{GL}^k(\nabla v) - \nabla I_{GL}^k v\|_{0,2,T,GL} \leq C \left(\frac{h_T}{k}\right)^{l-1} |v|_{l,2,T}. \tag{4.59}$$

Beweis

$$\begin{aligned}
\| \underbrace{I_{GL}^k(\nabla v) - \nabla I_{GL}^k v}_{\in \mathbb{Q}_k(T)} \|_{0,2,T,GL} &\underset{(4.44)}{\leq} C \|I_{GL}^k(\nabla v) - \nabla I_{GL}^k v\|_{0,2,T} \\
&\leq C \|\nabla(v - I_{GL}^k v)\|_{0,2,T} + C \|\nabla v - I_{GL}^k(\nabla v)\|_{0,2,T} \\
&\underset{(4.55)}{\leq} C \left(\frac{h_T}{k}\right)^{l-1} |v|_{l,2,T} + C \left(\frac{h_T}{k}\right)^{l-1} |\nabla v|_{l-1,2,T} \\
&\leq C \left(\frac{h_T}{k}\right)^{l-1} |v|_{l,2,T}. \quad \square
\end{aligned}$$

4.5 Inverse Ungleichungen

Ein wesentliches Instrument für die Analysis von numerischen Verfahren sind inverse Ungleichungen. Sie sind die Abschätzungen in den Normäquivalenzrelationen, die sich nicht direkt aus der Hölderschen Ungleichung ergeben.

Die Ungleichung nutzt die speziellen Eigenschaften der Tensorproduktdarstellung. Als Ergebnis liegt eine scharfe inverse Ungleichung vor, die mit Ausnahme der Differenz der Ableitungsordnungen und der räumlichen Dimension alle anderen Abhängigkeiten explizit enthält.

Die Abhängigkeit des Polynomgrads, die aus der *Nikolskii-Ungleichung* resultiert, lässt sich, wie das Beispiel

$$p(x) = \left(\frac{1 - T_n^2(x)}{1 - x^2} \right)^2$$

zeigt, nicht weiter verbessern (vgl. [Tim63, S. 236]), obwohl in [GNP08, Lemma 2.4] und [CB93, Lemma 1] anderslautende Resultate zu lesen sind. $T_n(x)$ entspricht hierbei dem n-ten *Tschebyscheff-Polynom*. Es handelt sich bei $p(x)$ um ein algebraisches Polynom, denn mit Hilfe der Darstellung der Tschebyscheff-Polynome durch

$$T_n(x) = \frac{1}{2} \left[(x + \sqrt{x^2 - 1})^n + (x - \sqrt{x^2 - 1})^n \right],$$

lassen sich für $1 - T_n^2(x)$ die Nullstellen -1 und 1 verifizieren.

Aufgrund der unterschiedlichen Aussagen bzgl. der Polynomabhängigkeit werden die Ungleichungen nach Nikolskii [Nik51] und Markov [HST37, Section III] auf Tensorprodukte verallgemeinert und in den Beweis der inversen Ungleichung nach [EG04, Lemma 1.138] integriert. Mit dieser Strategie kann das folgende Lemma gezeigt werden:

Lemma 4.5.1 (lokale inverse Ungleichung) *Für das Referenzelement $\{\hat{T}, \hat{P}, \hat{\Sigma}\}$ sei $l \geq 0$, so dass die Inklusion $\hat{P} \subset W^{l,\infty}(\hat{T})$ erfüllt ist. $\{\mathcal{T}_h^n\}_{h>0}$ sei eine lokal quasiuniforme Familie von affinen Partitionierungen mit $h < 1$. Ist zusätzlich $0 \leq m \leq l$, dann existiert*

1. *für $1 \leq p, q \leq \infty$ ein C abhängig von $l, m, p, q, d, \sigma_0, \hat{T}, P(\hat{T})$ mit*

$$\|v\|_{l,p,T} \leq C h_T^{m-l+d(\frac{1}{p}-\frac{1}{q})} \|v\|_{m,q,T} \quad \forall v \in P(T), \tag{4.60}$$

2. *für $1 \leq q \leq p \leq \infty$ und $\hat{T} = (-1, 1)^d$, $\hat{P} = \mathbb{Q}_k(\hat{T})$ ein $C(l, m, p, d, \sigma_0)$ mit*

$$\|v\|_{l,p,T} \leq C \left(\frac{h_T}{k^2} \right)^{m-l} \left(\frac{h_T}{2(q+1)k^2} \right)^{d(\frac{1}{p}-\frac{1}{q})} \|v\|_{m,q,T} \quad \forall v \in P(T). \tag{4.61}$$

Der Beweis dieses Lemmas wird nach der Präsentation der erwähnten Ungleichungen am Ende des Kapitels vorgelegt.

Lemma 4.5.2 (Nikolskii) *Für $0 < q \leq p \leq \infty$ und $\hat{v} \in \mathbb{Q}_k(\hat{T})$ gilt*

$$\|\hat{v}\|_{0,p,\hat{T}} \leq \left((q+1)k^2 \right)^{-d(1/p-1/q)} \|\hat{v}\|_{0,q,\hat{T}}. \tag{4.62}$$

Beweis Unter den angegebenen Voraussetzungen gilt nach [DL93, S. 102, Theorem 2.6] für $I = (-1, 1)$ die Ungleichung

$$\|\hat{v}\|_{0,p,I} \leq \left((q+1)k^2\right)^{-(1/p-1/q)} \|\hat{v}\|_{0,q,I}.$$

Damit gilt für $0 \leq i \leq d_x$ auch

$$\begin{aligned}&\|\hat{v}(\hat{x}_0, \ldots, \hat{x}_{i-1}, \cdot, \hat{x}_{i+1}, \ldots, \hat{x}_{d_x})\|_{0,\infty,I}\\ &\leq \left((q+1)k^2\right)^{1/q} \|\hat{v}(\hat{x}_0, \ldots, \hat{x}_{i-1}, \cdot, \hat{x}_{i+1}, \ldots, \hat{x}_{d_x})\|_{0,q,I}.\end{aligned}$$

Mit Hilfe dieser Ungleichung erfolgt der nächste Schritt

$$\begin{aligned}\|\hat{v}\|^q_{0,\infty,\hat{T}} &= \max_{\hat{x}_0 \in I} \cdots \max_{\hat{x}_{d_x} \in I} |\hat{v}(\hat{x}_0, \ldots, \hat{x}_{d_x})|^q\\ &\leq \max_{\hat{x}_0 \in I} \cdots \max_{\hat{x}_{d_x-1} \in I} ((q+1)k^2) \int_I |\hat{v}(\hat{x}_0, \ldots, \hat{x}_{d_x})|^q d\hat{x}_{d_x}\\ &\leq ((q+1)k^2) \max_{\hat{x}_0 \in I} \cdots \max_{\hat{x}_{d_x-2} \in I} \int_I \max_{\hat{x}_{d_x-1}} |\hat{v}(\hat{x}_0, \ldots, \hat{x}_{d_x})|^q \, d\hat{x}_{d_x}\\ &\leq ((q+1)k^2)^2 \max_{\hat{x}_0 \in I} \cdots \max_{\hat{x}_{d_x-2} \in I} \int_I \int_I |\hat{v}(\hat{x}_0, \ldots, \hat{x}_{d_x})|^q \, d\hat{x}_{d_x-1} d\hat{x}_{d_x}\\ &\leq \cdots \leq ((q+1)k^2)^d \|\hat{v}\|^q_{0,q,\hat{T}},\ 0 < q < \infty.\end{aligned}$$

Die Anwendung der Interpolationsungleichung (2.28) liefert mit obiger Ungleichung die Behauptung:

$$\begin{aligned}\|\hat{v}\|_{0,p,\hat{T}} &\leq \|\hat{v}\|^{\frac{q}{p}}_{0,q,\hat{T}} \|\hat{v}\|^{1-\frac{q}{p}}_{0,\infty,\hat{T}}\\ &\leq \|\hat{v}\|^{\frac{q}{p}}_{0,q,\hat{T}} \left((q+1)k^2\right)^{\frac{d}{q}(1-\frac{q}{p})} \|\hat{v}\|^{1-\frac{q}{p}}_{0,q,\hat{T}}\\ &\leq \left((q+1)k^2\right)^{-d(\frac{1}{p}-\frac{1}{q})} \|\hat{v}\|_{0,q,\hat{T}}, \quad 0 < q \leq p \leq \infty.\end{aligned}$$ □

Lemma 4.5.3 (Verallgemeinerte *Markov-Ungleichung*) *Für $v \in \mathbb{Q}_k(I)$, $I = (-1, 1)$ und $p > 1$ gilt*

$$\|v'\|_{0,p,I} \leq C_M(p) k^2 \|v\|_{0,p,I}, \tag{4.63}$$

mit

$$C_{M,p} = C_M(p) = 2(p-1)^{1/p-1} \left(p + \frac{1}{k}\right) \left(1 + \frac{p}{kp - p + 1}\right)^{k-1+1/p}.$$

Für $p = \infty$ gilt sogar $C_{M,\infty} = 1 < \lim_{p\to\infty} C_M(p) = 2e$. Darüber hinaus gilt für alle $p \in \mathbb{N}$: $C_M(p) \leq C_M = 6e^{1+1/e}$.

Beweis In [HST37, Section III] wird für $p < \infty$ die Behauptung bewiesen. Für $p \in \mathbb{N}$: $C_M(p) \leq C_M = 6e^{1+1/e}$ siehe [MMR94, S. 590] und für $p = \infty$ ergibt sich das Ergebnis mit obiger Substitution und [DL93, S. 98, Theorem 1.4]. □

Bemerkung 4.5.4 Es ist möglich, die Konstanten $C_M(p)$ weiter zu verbessern. So zeigt [Bar98, Corollary 2.10] die Existenz eines $\tilde{C}_M(p)$ mit $\lim_{p\to\infty} \tilde{C}_M(p) = 1$, $p > 2$.

Folgerung 1 *Für* $1 \leq p \leq \infty$ *und* $\hat{v} \in \mathbb{Q}_k(\hat{T})$ *gilt*

$$\|\hat{v}\|_{l,p,\hat{T}} \leq \binom{d+l}{l}^{\frac{1}{p}} \left(C_{M,p}k^2\right)^l \|\hat{v}\|_{0,p,\hat{T}} \tag{4.64}$$

mit den Konstanten von (4.63).

Beweis Mit Iteration über die Ableitungsordnung genügt es für

$$\|\partial^\alpha \hat{v}\|_{0,p,\hat{T}} \leq (C_{M,p}k^2)^{|\alpha|} \|\hat{v}\|_{0,p,\hat{T}}$$

zu zeigen, dass die Ungleichung für $|\alpha| = 1$ gilt. Sei dazu $\alpha = e_i$, mit e_i als der i-te Einheitsvektor.

Für $p < \infty$ gilt mit der verallgemeinerten Markov-Ungleichung für $0 \leq i \leq d_x$:

$$\begin{aligned}&\|\hat{v}'(\hat{x}_0, \ldots, \hat{x}_{i-1}, \cdot, \hat{x}_{i+1}, \ldots, \hat{x}_{d_x})\|^p_{0,p,I} \\ &\leq C^p_{M,p}k^{2p}\|\hat{v}(\hat{x}_0, \ldots, \hat{x}_{i-1}, \cdot, \hat{x}_{i+1}, \ldots, \hat{x}_{d_x})\|^p_{0,p,I}\end{aligned}$$

und somit für das Referenzelement

$$\begin{aligned}\|\partial^\alpha \hat{v}\|^p_{0,p,\hat{T}} &= \int_I \ldots \int_I |\partial_i \hat{v}(\hat{x})|^p \, d\hat{x}_i d\hat{x}_0 \ldots d\hat{x}_{i-1} d\hat{x}_{i+1} \ldots d\hat{x}_{d_x} \\ &\leq C^p_{M,p}k^{2p}\|\hat{v}\|^p_{0,p,\hat{T}},\end{aligned}$$

während für $p = \infty$ ein $\hat{y} \in \hat{T}$ existiert, so dass

$$\begin{aligned}\|\partial^\alpha \hat{v}\|_{0,\infty,\hat{T}} = |\partial_i \hat{v}(\hat{y})| &= \max_{\hat{x}_i \in I} |\partial_i \hat{v}(\hat{y}_0, \ldots, \hat{y}_{i-1}, \hat{x}_i, \hat{y}_{i+1}, \ldots, \hat{y}_{d_x})| \\ &\leq C_{M,\infty}k^2\|\hat{v}(\hat{y}_0, \ldots, \hat{y}_{i-1}, \cdot, \hat{y}_{i+1}, \ldots, \hat{y}_{d_x})\|_{0,\infty,I} \\ &\leq C_{M,\infty}k^2\|\hat{v}\|_{0,\infty,\hat{T}}\end{aligned}$$

gilt. Ausgehend von diesen Betrachtungen resultieren

$$\begin{aligned}\|\hat{v}\|^p_{l,p,\hat{T}} &= \sum_{i=0}^{l} \sum_{|\alpha|=i} \|\partial^\alpha \hat{v}\|^p_{0,p,\hat{T}} \leq \sum_{i=0}^{l} \sum_{|\alpha|=i} \left(C_{M,p}k^2\right)^{|\alpha|p} \|\hat{v}\|^p_{0,p,\hat{T}} \\ &= \sum_{i=0}^{l} \binom{d+i-1}{i} \left(C_{M,p}k^2\right)^{ip} \|\hat{v}\|^p_{0,p,\hat{T}} \\ &\leq \binom{d+l}{l} \left(C_{M,p}k^2\right)^{lp} \|\hat{v}\|^p_{0,p,\hat{T}}\end{aligned}$$

und

$$\begin{aligned}\|\hat{v}\|_{l,\infty,\hat{T}} &= \max_{0 \leq |\alpha| \leq l} \|\partial^\alpha \hat{v}\|_{0,\infty,\hat{T}} \leq \max_{0 \leq |\alpha| \leq l} \left(C_{M,\infty}k^2\right)^{|\alpha|} \|\hat{v}\|_{0,\infty,\hat{T}} \\ &\leq \left(C_{M,\infty}k^2\right)^l \|\hat{v}\|_{0,\infty,\hat{T}}.\end{aligned}$$

□

Der noch fehlende Beweis der lokalen inversen Ungleichung aus Lemma 4.5.1 kann nun wie folgt geführt werden.

Beweis (von Lemma 4.5.1) Der erste Fall wird in [EG04, Lemma 1.138] behandelt. Im Spezialfall des Lagrangeschen Referenzelements bleibt die Beweisidee erhalten, wobei allerdings die Lemmata und Folgerungen aus diesem Kapitel verwendet werden. Wie üblich zeigt man die Aussage zuerst für $m = 0$. Mit (4.64) und (4.62) folgt zunächst für das Referenzelement

$$\|\hat{v}\|_{l,p,\hat{T}} \leq \binom{d+l}{l}^{\frac{1}{p}} \left(C_{M,p}k^2\right)^l \left((q+1)k^2\right)^{-d(\frac{1}{p}-\frac{1}{q})} \|\hat{v}\|_{0,q,\hat{T}}. \tag{4.65}$$

Um diese Abschätzung für ein transformiertes Element zu erhalten, verwende (4.6), (4.8) und (4.9) für $0 \leq j \leq l$:

$$\begin{aligned}
|v|^p_{j,p,T} &\leq \left\{C_{j,d}\|J_T^{-1}\|^j_{l^2}|det(J_T)|^{1/p}\right\}^p |\hat{v}|^p_{j,p,\hat{T}} \\
&\leq \left\{C_{j,d}\left(2\sqrt{d}\sigma_0 h_T^{-1}\right)^j \left(\frac{h_T}{2}\right)^{\frac{d}{p}}\right\}^p \|\hat{v}\|^p_{j,p,\hat{T}} \\
&\underset{(4.65)}{\leq} \left\{\binom{d+j}{j}^{\frac{1}{p}} \tilde{C}(j,p,d,\sigma_0)\left(\frac{h_T}{k^2}\right)^{-j}\left(\frac{h_T}{2}\right)^{\frac{d}{p}}\left((q+1)k^2\right)^{-d(\frac{1}{p}-\frac{1}{q})}\right\}^p \|\hat{v}\|^p_{0,q,\hat{T}}
\end{aligned}$$

mit $\tilde{C}(j,p,d,\sigma_0) = C_{j,d}\left(2\sqrt{d}\sigma_0 C_{M,p}\right)^j$. Die Rücktransformation erfolgt mit (4.5), $l = 0$ und liefert

$$|v|_{j,p,T} \leq \binom{d+j}{j}^{\frac{1}{p}} \tilde{C}(j,p,d,\sigma_0)\left(\frac{h_T}{k^2}\right)^{-j}\left(\frac{h_T}{2(q+1)k^2}\right)^{d(\frac{1}{p}-\frac{1}{q})} \|v\|_{0,q,T}. \tag{4.66}$$

Für die entsprechende Sobolev-Norm gilt somit unter der Voraussetzung $h_T^{-1} \geq 1$ für $0 \leq j \leq l$

$$\begin{aligned}
\|v\|^p_{j,p,T} &= \sum_{s=0}^{j} |v|^p_{s,p,T} \\
&\leq \sum_{s=0}^{j} \binom{d+s}{s}\left\{\tilde{C}(s,p,d,\sigma_0)\left(\frac{h_T}{k^2}\right)^{-s}\left(\frac{h_T}{2(q+1)k^2}\right)^{d(\frac{1}{p}-\frac{1}{q})}\right\}^p \|v\|^p_{0,q,T} \qquad (4.67) \\
&\leq \binom{d+1+j}{j}\left\{\tilde{C}(j,p,d,\sigma_0)\left(\frac{h_T}{k^2}\right)^{-j}\left(\frac{h_T}{2(q+1)k^2}\right)^{d(\frac{1}{p}-\frac{1}{q})}\right\}^p \|v\|^p_{0,q,T}
\end{aligned}$$

und die Behauptung ist für $m = 0$ bewiesen. Sei jetzt $0 \leq m \leq l$ und α ein Multiindex mit $0 \leq |\alpha| \leq l$. Im Fall von $|\alpha| \leq l - m$ folgt mit (4.67) aus

$$\|\partial^\alpha v\|^p_{0,p,T} \leq \sum_{|\alpha|\leq l-m} \|\partial^\alpha v\|^p_{0,p,T} = \|v\|^p_{l-m,p,T}$$

die Ungleichung

$$\begin{aligned}
\|\partial^\alpha v\|_{0,p,T} &\leq C(l-m,p,d,\sigma_0)\left(\frac{h_T}{k^2}\right)^{m-l}\left(\frac{h_T}{2(q+1)k^2}\right)^{d(\frac{1}{p}-\frac{1}{q})} \|v\|_{0,q,T} \\
&\leq C(l-m,p,d,\sigma_0)\left(\frac{h_T}{k^2}\right)^{m-l}\left(\frac{h_T}{2(q+1)k^2}\right)^{d(\frac{1}{p}-\frac{1}{q})} \|v\|_{m,q,T}
\end{aligned}$$

mit $C(j,p,d,\sigma_0) = \binom{d+1+j}{j}^{1/p}\tilde{C}(j,p,d,\sigma_0)$. Diese Ungleichung ist auch für $l-m \leq |\alpha| \leq l$ richtig, denn hier lassen sich zwei weitere Multiindizes β und γ angeben, für die $\alpha = \beta + \gamma$, $|\beta| = l-m$ und $|\gamma| \leq m$ erfüllt ist. Analog zum ersten Fall folgt jetzt

$$\begin{aligned}\|\partial^\alpha v\|_{0,p,T} &= \|\partial^\beta(\partial^\gamma v)\|_{0,p,T} \\ &\leq C(l-m,p,d,\sigma_0)\left(\frac{h_T}{k^2}\right)^{m-l}\left(\frac{h_T}{2(q+1)k^2}\right)^{d(\frac{1}{p}-\frac{1}{q})}\|\partial^\gamma v\|_{0,q,T} \\ &\leq C(l-m,p,d,\sigma_0)\left(\frac{h_T}{k^2}\right)^{m-l}\left(\frac{h_T}{2(q+1)k^2}\right)^{d(\frac{1}{p}-\frac{1}{q})}\|v\|_{m,q,T}.\end{aligned}$$

Da diese Ungleichung für $0 \leq |\alpha| \leq l$ gilt, kann auch $\|v\|^p_{l,p,T}$ über die Definition abgeschätzt werden und man erhält die Behauptung mit

$$C(l,m,p,d,\sigma_0) = C_{l-m,d}\left(2\sqrt{d}\sigma_0 C_{M,p}\right)^{l-m}\binom{d+1+l-m}{l-m}^{1/p}\binom{d+l}{l}^{1/p}.$$

Die Gleichheit tritt bei der inversen Ungleichung ein, falls $l = m = 0$ und $p = q$ ist. □

Lemma 4.5.5 (globale inverse Ungleichung) *Unter den Voraussetzungen des Lemmas* 4.5.1 *existiert für eine quasiuniforme Familie* $\{\mathcal{T}_h^n\}$ *von Partitionierungen für alle* $v \in W_h$ *und*

1. *für* $1 \leq p,q \leq \infty$ *ein* C *abhängig von* $l,m,p,q,d,\sigma_0,\hat{T},P(\hat{T}),C_{qu}$ *mit*

$$\left(\sum_{T\in\mathcal{T}_h^n}\|v\|^p_{l,p,T}\right)^{\frac{1}{p}} \leq Ch^{m-l+min(0,d(\frac{1}{p}-\frac{1}{q}))}\left(\sum_{T\in\mathcal{T}_h^n}\|v\|^q_{m,q,T}\right)^{\frac{1}{q}}, \tag{4.68}$$

2. *für* $1 \leq q \leq p \leq \infty$ *und* $\hat{P} = \mathbb{Q}_k(\hat{T})$ *ein* $C = C(l,m,p,d,\sigma_0)$ *mit*

$$\left(\sum_{T\in\mathcal{T}_h^n}\|v\|^p_{l,p,T}\right)^{\frac{1}{p}} \leq C\left(\frac{C_{qu}h}{k^2}\right)^{m-l}\left(\frac{C_{qu}h}{2(q+1)k^2}\right)^{d(\frac{1}{p}-\frac{1}{q})}\left(\sum_{T\in\mathcal{T}_h^n}\|v\|^q_{m,q,T}\right)^{\frac{1}{q}}. \tag{4.69}$$

Beweis Der Beweis der ersten Abschätzung ist in [EG04, Lemma 1.141] zu finden. Die zweite Abschätzung folgt mit

$$C = C(l,m,p,d,\sigma_0)$$

aus Lemma 4.5.1 und der Quasiuniformität von $\{\mathcal{T}_h^n\}_{h>0}$

$$\sum_{T\in\mathcal{T}_h^n}\|v\|^p_{l,p,T} \leq C^p\left(\frac{C_{qu}h}{k^2}\right)^{p(m-l)}\left(\frac{C_{qu}h}{2(q+1)k^2}\right)^{pd(\frac{1}{p}-\frac{1}{q})}\sum_{T\in\mathcal{T}_h^n}\|v\|^p_{m,q,T}. \tag{4.70}$$

Das Ziehen der p-ten Wurzel liefert zusammen mit $\|\cdot\|_{l^p} \leq \|\cdot\|_{l^q}$ die Behauptung für $p,q \neq \infty$. Der Beweis für $p,q = \infty$ erfolgt entsprechend aus den zugehörigen Normen. □

5 Discontinuous-Galerkin Approximation

5.1 Formulierung der Methode

Dieses Kapitel beschreibt die Mechanismen zur Lösung hyperbolischer, partieller Differentialgleichungen erster Ordnung. Die erste Discontinuous-Galerkin-Methode für diese Aufgabe wurde in [RH73] vorgestellt. Numerische Tests, Stabilitätsaussagen und a priori Abschätzungen z.B. aus [JP86, Theoreme 2.1, 4.3], [EG04, Theoreme 5.73, 6.56] und [Joh87, S. 194, 208] zeigen für stationäre und instationäre Probleme sehr deutlich die Überlegenheit gegenüber der Standard-Galerkin-Methode.

Sei x ein Punkt auf $\tau = \overline{\partial T^+} \cap \overline{\partial T^-}$, wobei $T^+ = T$ und T^- ein Nachbarelement bezeichnet. $n_T = n_T^+$ und n_T^- bezeichnen die dazugehörigen Einheitsnormalenvektoren zu $\partial T^\pm$ am Punkt x und $n = n^+$, n^- den Einheitsnormalenvektor bzgl. $\tau \in R_{n,n+1}$ bzw. $\tau \in \Lambda_{n,n+1}$. Damit lassen sich folgende Definitionen vornehmen (vgl. Abbildung 4.1):

$$v^\pm(x) = \lim_{\mu \to +0} v(x - \mu n^\pm), \quad v_\pm^n(x) = v(t_n \pm 0, x_1, \ldots, x_d). \tag{5.1}$$

Es gilt $U^-(x) = g_D(x)$ falls $x \in \Sigma_T$. Im Folgenden sei $f(v) = (f_0(v), f_1(v), \ldots, f_{d_x}(v))^T$ und $f_x(v) = (f_1(v), \ldots, f_{d_x}(v))^T$ mit $f_0(v) = v$. Eine schwache Formulierung von (3.12), (3.13) und (3.14) ist z.B.:

Finde $u \in W = W^{1,\infty}(Q_T)$ *derart, dass*

$$a(u, v) = 0 \quad \forall v \in W^{1,2}(Q_T, \mathcal{T}_h) \tag{5.2}$$

erfüllt ist.

Hierbei ist

$$a(v, w) = \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \left\{ \int_T L(v) w \, dx + \int_{\partial T} (H(v) - f(v_T^+) \cdot n_T^+) w_T^+ \, ds \right\} \tag{5.3}$$

mit der *modifizierten Flussdichte von Lax-Friedrich*

$$H(U) = \frac{1}{2} \left(f(U_T^+) + f(U_T^-) \right) \cdot n_T^+ + C_T \left(U_T^+ - U_T^- \right). \tag{5.4}$$

Die allg. *Flussdichte von Lax-Friedrich* lautet für $f : \mathbb{R}^m \to \mathbb{R}^{d \times m}$, $f_0(v) = v, m \in \mathbb{N}$

$$H(U) = \frac{1}{2} \left(f(U_T^+) + f(U_T^-) \right) \cdot n_T^+ + \frac{1}{2} \alpha(U_T^+, U_T^-) \left(U_T^+ - U_T^- \right), \tag{5.5}$$

wobei $\alpha(v,w) = \max\limits_{1\le i\le m}\sup\limits_{z\in[v,w]}\left\{|\lambda_i(f'(z)\cdot n_T^+)|\right\}$ mit dem Eigenwert $\lambda_i(A)$ von A ist (vgl. [BO04, Section 4.4]).
In der allgemeinen Flussdichte von Lax-Friedrich und $m = 1$ wäre somit

$$C_T(x,v,w) = \begin{cases} \frac{1}{2} & n_T^+(x) = \pm(1,0,\dots,0), \\ \frac{1}{2}\sup\limits_{z\in[v(x),w(x)]} |{f_x}'(z)\cdot {n_x}^+(x)| & \text{sonst.} \end{cases} \tag{5.6}$$

Die modifizierte Flussdichte von Lax-Friedrich hingegen entsteht aus der Definition

$$C_T(x) = \begin{cases} \frac{1}{2} & n_T^+(x) = \pm(1,0,\dots,0), \\ C_0^{\partial\Omega} & x\in\partial\Omega, \\ C_0^{\Omega} & \text{sonst,} \end{cases} \tag{5.7}$$

mit $C_0^{\partial\Omega}, C_0^{\Omega} > 0$ konstant. Zusätzlich wird

$$\begin{aligned} \|f'\|_{0,\infty,\cdot|\partial\Omega} &= \max_{v|\partial\Omega}\|f'(v)\|_{l^2} \le \frac{1}{2}C_0^{\partial\Omega}, \\ \|f'\|_{0,\infty,\mathbb{R}} &= \max_{r\in\mathbb{R}}\|f'(r)\|_{l^2} \le C_0^{\Omega} \end{aligned} \tag{5.8}$$

und

$$f(0) = 0 \tag{5.9}$$

gefordert.

Bemerkung 5.1.1 Eine weitere Möglichkeit, die Aufgabe zu formulieren, liefert die partielle Integration von (5.3) :

$$a(v,w) = \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\left\{-\int_T f(v)\cdot\nabla w\,dx + \int_{\partial T} H(v)w_T^+\,ds\right\}. \tag{5.10}$$

Ist $f(v) = bv$ linear mit einem Vektorfeld $b\in C(\overline{Q}_T)^d$, $b_0 = 1$ und $\nabla\cdot b\in L^\infty(Q_T)$, so folgt aus (5.2) mit $L(u) = b\cdot\nabla u$ und (5.6) direkt die Discontinuous-Galerkin-Methode

$$\begin{aligned} &\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\int_T (b\cdot\nabla U)\,v\,dx + \sum_{n=1}^{N-1}\int_\Omega (U_+^n - U_-^n)v_+^n\,dx_x + \int_\Omega U_+^0 v_+^0\,dx_x \\ &+ \sum_{n=0}^{N-1}\int_{R_{n,n+1}^i} |b_x\cdot {n_x}^+|(U^+ - U^-)v^+\,ds + \sum_{n=0}^{N-1}\int_{\Lambda_{n,n+1}^-} |b_x\cdot {n_x}^+|U^+v^+\,ds \\ &= \int_\Omega u_0 v_+^0\,dx_x + \sum_{n=0}^{N-1}\int_{\Lambda_{n,n+1}^-} |b_x\cdot {n_x}^+|g_D v^+\,ds \quad \forall v\in W_h^n, \end{aligned} \tag{5.11}$$

bzw. in stationären Fall $U_+^n - U_-^n = 0$ für alle $0\le n\le N-1$

$$\sum_{T\in\mathcal{T}_h^n}\Big\{\int_T (b_x\cdot\nabla U)\,v\,dx + \int_{\partial T^-} |b_x\cdot {n_x}^+|(U_T^+ - U_T^-)v_T^+\,ds = 0. \tag{5.12}$$

Dies ist die Formulierung aus [JP86], jedoch mit $v_{JP}^{\pm}(x) = \lim_{\mu \to \pm 0} v(x + \mu b_x)$. Auf ∂T^- gilt dann $v_{JP}^{\pm}(x) = \lim_{\mu \to \pm 0} v(x + \mu b_x n \cdot n) = \lim_{\mu \to \pm 0} v(x - \mu n) = v^{\pm}(x)$.

Für diese Methode ist es möglich einen Fehler von der Ordnung $\mathcal{O}(h^{k+1/2})$ bzgl. einer Norm, die auch die Ableitung in Stromlinienrichtung beinhaltet, zu zeigen. Diese Eigenschaft weist die *Standard-Galerkin-Methode* nur in Verbindung mit einer Stabilisierungsmethode auf. Numerische Tests zeigen zudem, dass die Methode bei Partitionierungen aus der Praxis häufig einen Fehler der Form $\mathcal{O}(h^{k+1})$ besitzt. Trotz dieser erweiterten Kontrolle des Fehlers treten bei moderaten Gitterweiten nichtphysikalische Oszillationen auf.

5.2 Shock-capturing

Als Ausweg bietet sich das Hinzufügen eines Termes mit *isotroper künstlicher Diffusion*, der für $h \to 0$ verschwindet, an. Eine solche Methode, die z.B. den klassischen künstlichen Diffusionsterm $Ch_T(\nabla U, \nabla v)_{0,2,T}$ enthält, erzeugt nichtoszillierende Lösungen. Der Nachteil dieses Vorgehens liegt zum einen in der Addition von Diffusion in sämtliche Richtungen und zum anderen in der Skalierung der Diffusion. Als Konsequenz ist der Fehler bestenfalls von der Ordnung $\mathcal{O}(h)$. Sogar bei einer glatten, kontinuierlichen Lösung im Zusammenhang mit einer Diskretisierung höherer Ordnung kann die Fehlerordnung nicht verbessert werden. Die Aufgabe besteht nun darin, nicht zu viel künstliche Diffusion in der richtigen Art und Weise hinzuzufügen, so dass der Fehler von akzeptabler Ordnung ist. Im Gegenzug muss jedoch auch sichergestellt werden, dass die zusätzliche Diffusion ausreichend ist, um übermäßiges Oszillieren zu vermeiden.

Eine Möglichkeit, die zusätzliche Diffusion zu konstruieren, besteht in der Verwendung einer Diskretisierungsmethode, die mit einer Konstanten $C > 0$ unabhängig von h eine Ungleichung der Form

$$\|U\|_{0,\infty,Q_T} \leq C \left\{ \|u_0\|_{0,\infty,\Omega} + \|g_D\|_{0,\infty,\Sigma_T} \right\}, \quad \forall h > 0$$

erfüllt. Diese Aussage garantiert, unabhängig von der Gitterweite h, zusätzliche Stabilität der diskreten Lösung.

Eine solche Methode im Discontinuous-Galerkin Kontext wird in [JJS95] vorgestellt. Zum Ziel führte hier die Addition von *anisotroper künstlicher Diffusion* in Stromlinienrichtung und eine residual basierte, isotrope Diffusion.

Alternative Shock-capturing-Methoden finden sich in den Übersichten [JK07] und [JK08b]. Die isotrope Diffusion aus [JJS95] findet sich hier in einer normierten Variante unter [JK07, (18)] wieder.

In dieser Arbeit hingegen soll u.a. untersucht werden, ob die genannte Stabilitätsaussage auch mit einer isotropen Diffusion garantiert werden kann. Die dafür entwickelte Analysis (siehe Lemma 5.4.8) verallgemeinert darüber hinaus für quasiuniforme Familien von Partitionierungen $\{\mathcal{T}_h^n\}_{h>0}$ die $L^\infty(L^\infty)$-Abschätzung aus [Sze89a], [Sze91] und [JJS95] auf Ansatzfunktionen höherer Ordnung.

Ferner soll die Frage beantwortet werden, ob die Vernachlässigung des Stromliniendiffusionsterms der Methode erlaubt mit einem Fehler aufzuwarten, dessen Ordnung nicht durch Eins

beschränkt ist. Als Vorteil gegenüber der Methode in [JJS95] ist die geringere Abhängigkeit vom Differentialoperator L zu sehen.

Die oben angesprochene Art der Stabilisierung wird realisiert, indem die künstliche Diffusion geeignet auf die feinen Skalen projiziert wird. Dies ergibt einen Stabilisierungsterm, der Fluktuationen des Gradienten enthält. Eine derartige Vorgehensweise wird als *lokale Projektions Stabilisierung* (LPS) bezeichnet [BB04], [Hei07], [MST07], [KL09]. Im Gegensatz dazu existieren in der Literatur Methoden, die die feinen Skalen mit künstlicher Diffusion stabilisieren, d.h. es wird der Gradient der Fluktuationen verwendet [Gue99], [EG04]. Beide Techniken können auch als ein Modell für den Einfluss der nichtaufgelösten Skalen auf die feinen Skalen gedeutet werden und sind daher ebenso den *Variationellen Mehrskalen (VMS) Methoden* zuzuordnen. Siehe hierzu unter anderem die Arbeiten [JK08a], [JKL06] und [KR05], inklusive der dort angegebenen Quellen.

In der Literatur werden die groben Skalen zum einen durch ein Gitter mit größerer Maschenweite (*subgrid artificial viscosity*) und zum anderen durch einen Polynomraum niedrigerer Ordnung dargestellt. Im weiteren Verlauf werden die groben Skalen allerdings durch eine Linearkombination von ausgewählten Basispolynomen aus $\mathbb{Q}_k(T)$ repräsentiert. Die Verwirklichung dieses Vorhabens leistet der schon in (4.50) definierte Fluktuationsoperator P_h' bzw. die „VMS-Skalierung“ der künstlichen Diffusion $\hat{\epsilon}_{\text{vms}}(U,v)$ aus (4.54):

$$(P_h'(\nabla U), P_h'(\nabla v))_{0,T,h} = \hat{\epsilon}_{\text{vms}}(U,v)\,(\nabla U, \nabla v)_{0,T} \approx \underbrace{s(\hat{\epsilon}_{\text{vms}}(U,v))}_{=\hat{\epsilon}^{\text{coer}}_{\text{vms}}(U,v)}\,(\nabla U, \nabla v)_{0,T}\,.$$

Die Funktion $s : \mathbb{R} \to \mathbb{R}$ ist eine Lipschitz-stetige Funktion, die im Wesentlichen wie die Identität wirkt. Allerdings ist sie in einer Umgebung der Null modifiziert, so dass für nichtnegatives $\hat{\epsilon}_{\text{vms}}(U,v)$ die für die Analysis notwendige Bedingung $\hat{\epsilon}^{\text{coer}}_{\text{vms}}(U,v) \geq \hat{\epsilon}^{\text{min}}_{\text{vms}} > 0$ erfüllt ist. Mit der Lipschitz-Stetigkeit von s folgt direkt

$$|s(\frac{x}{z})z - s(\frac{y}{z})z| \leq \mathcal{L}_{[s]}|x-y| \quad \forall z \in \mathbb{R}. \tag{5.13}$$

Eine Möglichkeit besteht in der Wahl von

$$s(x) = C_4 + (x - C_4)e^{-\frac{7\cdot 3^8 C_4^8}{8\cdot 2^8 (x-C_4)^8}}, C_4 > 0. \tag{5.14}$$

Bei dieser Wahl liegt die einzige Nullstelle zwischen $-C_4$ und $-C_4/2$. Ferner ist

$$s(x) \geq \hat{\epsilon}^{\text{min}}_{\text{vms}} \approx C_4,\ x \geq 0$$

und an den beiden Wendestellen $-C_4/2$ und $5/2C_4$ nimmt $s'(x)$ das Maximum $8\exp(-7/8)$ an. Für kleines $\hat{\epsilon}_{\text{vms}}(U,v) > 0$ erlischt somit der VMS-Charakter des Verfahrens auf dem entsprechenden Element T und resultiert schlechtestenfalls in einem Fehler der Ordnung $\mathcal{O}(h^{1/2}/k^{1/2})$ (vgl. Theorem 6.1.3).

Betrachte jetzt das Problem:

Finde $U \in W_h$ derart, dass für $n = 0, 1, \ldots, N-1$ und $U \equiv U_{Q_{n,n+1}} \in W_h^n$

$$a(U,v) + \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}(U)\hat{\epsilon}^{\text{coer}}_{\text{vms}}(U,v)\,(\nabla U, \nabla v)_{0,T} = 0 \quad \forall v \in W_h^n. \tag{5.15}$$

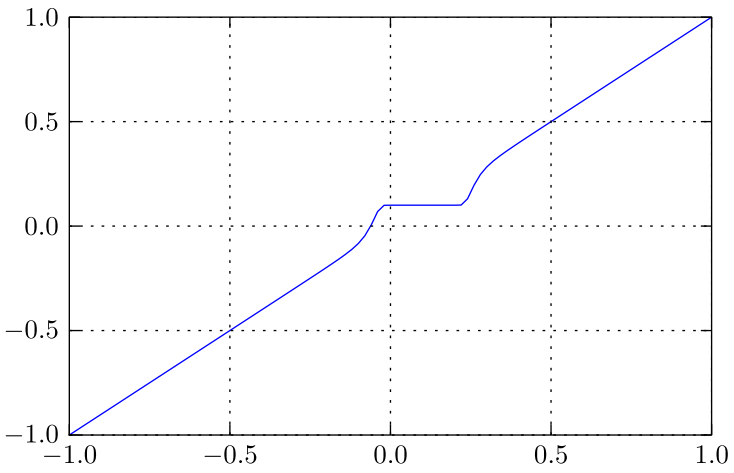


Abbildung 5.1: $s(x) = C_4 + (x - C_4)e^{-\frac{7 \cdot 3^8 C_4^8}{8 \cdot 2^8 (x - C_4)^8}}$, $C_4 = 0.1$

In dieser Formulierung ist

$$\begin{aligned}\hat{\epsilon}(U) &= \hat{\epsilon}_1(U) + \hat{\epsilon}_2(U) + \hat{\epsilon}_3(U) \\ &= C_1(k) h_T \max_T \left(\|f'(U)\|_{l^2}\right) + C_2(k) h_T^{2-\beta} \max_T(\|f'(U)\|_{l^2}) \max_T(\|\nabla U\|_{l^2}) \\ &+ C_3(k) h_T^{2-\beta} R(U), \quad 0 < \beta < 1/2\end{aligned} \tag{5.16}$$

mit

$$R(U) = h_T^{-1} \left(\max_{\partial^* T} \left(\left| \left[f(U_T^+) - f(U_T^-) \right] \cdot n_T^+ \right| \right) + \max_{\partial^* T} \left(C_T \left| U_T^+ - U_T^- \right| \right) \right) \tag{5.17}$$

und $\partial^* T = \{x \in \partial T : x \notin R_{n+1}\}$ konstant auf T. Bei der Konvergenzanalysis wird darüber hinaus die Abhängigkeit des Fehlers vom Polynomgrad betrachtet. Dazu wird bei der künstlichen Diffusion das Innenprodukt $(\cdot, \cdot)_{0,T,GL}$ verwendet und die Definition des Koeffizienten $\hat{\epsilon}(U)$ modifiziert:

$$\begin{aligned}\hat{\epsilon}(U) &= \hat{\epsilon}_1(U) + \hat{\epsilon}_2(U) + \hat{\epsilon}_3(U) \\ &= C_1 \frac{h_T}{k} \max_T \left(\|f'(U)\|_{l^2}\right) + C_2 \left(\frac{h_T}{k^{2+d/2}} \right)^{2-\beta} \max_T(\|f'(U)\|_{l^2}) \max_T(\|\nabla U\|_{l^2}) \\ &+ C_3 \left(\frac{h_T}{k^{2+d/2}} \right)^{2-\beta} R(U)\end{aligned} \tag{5.18}$$

mit $0 < \beta < \min\left(\frac{1}{2}, \frac{3}{d+4}\right)$ und

$$R(U) = \frac{k}{h_T} \left(\max_{\partial^* T} \left(\left| \left[f(U_T^+) - f(U_T^-) \right] \cdot n_T^+ \right| \right) + \max_{\partial^* T} \left(C_T \left| U_T^+ - U_T^- \right| \right) \right). \tag{5.19}$$

Die Konstanten $C_1, C_2, C_3 > 0$ werden jeweils in den entsprechenden Kapiteln definiert. Neben diesen Konstanten sind β, die verwendete Quadraturformel und der Raum mit den groben Skalen $P_h^{K,k}(\nabla \mathbb{Q}_k(T)) \subseteq \nabla \mathbb{Q}_k(T)$ Parameter des künstlichen Diffusionsterms. Für $P_h^{K,k}(\nabla \mathbb{Q}_k(T))$ sind theoretisch die beiden Grenzfälle

1. $P_h^{K,k}(\nabla \mathbb{Q}_k(T)) = \nabla \mathbb{Q}_k(T)$: Der Grobraum enthält alle Skalen, so dass keine feinskaligen Elemente auftreten und der künstliche Diffusionsterm verschwindet.
2. $P_h^{K,k}(\nabla \mathbb{Q}_k(T)) = \{0\}$: Die Stabilisierung wirkt auf alle Skalen.

möglich, wobei für den ersten Fall nichts zu zeigen ist.

Bemerkung 5.2.1 Zur Existenz und Eindeutigkeit der diskreten Lösung vgl. z.B. [JS86b].

Bemerkung 5.2.2 Die schwache Formulierung (5.15) ergibt für $k = 0$ auch eine *Finite-Volumen-Methode* nach [BO04, Definition 2.3]. Vgl. dazu auch den Abschnitt 2.2.6 in der gleichen Quelle.

Bemerkung 5.2.3 Unter den Voraussetzungen von (5.11) und mit der Definition (5.7) ergibt sich die Formulierung

$$\begin{aligned}
&\sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \left\{ \int_T (b \cdot \nabla U)\, v \, dx + \hat{\epsilon}(U) \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, v) \left(\nabla U, \nabla v\right)_{0,T} \right\} \\
&+ \sum_{n=1}^{N-1} \int_\Omega (U_+^n - U_-^n) v_+^n \, dx_x + \int_\Omega U_+^0 v_+^0 \, dx_x \\
&+ \sum_{n=0}^{N-1} \int_{R_{n,n+1}^i} \frac{1}{2} b_x \cdot {n_x}^+ (U^- - U^+)(v^+ + v^-) + C_0^\Omega (U^+ - U^-)(v^+ - v^-) \, ds \\
&+ \sum_{n=0}^{N-1} \int_{\Lambda_{n,n+1}^-} -\frac{1}{2} b_x \cdot {n_x}^+ U^+ v^+ + C_0^{\partial\Omega} U^+ v^+ \, ds \\
&= \int_\Omega u_0 v_+^0 \, dx_x + \sum_{n=0}^{N-1} \int_{\Lambda_{n,n+1}^-} -\frac{1}{2} b_x \cdot {n_x}^+ g_D v^+ + C_0^{\partial\Omega} g_D v^+ \, ds \quad \forall v \in W_h^n.
\end{aligned} \tag{5.20}$$

5.3 $L^\infty(L^2)$-Abschätzung der diskreten Lösung

Dieser Abschnitt dient dazu einige Bezeichnungen und Hilfsaussagen für die $L^\infty(L^\infty)$-Abschätzung einzuführen und mit dem Theorem 5.3.1 das Hauptresultat dieses Kapitels vorzubereiten.

Für die Definition

$$\begin{aligned}
b(v, w) = \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \Big\{ &\int_T \nabla \cdot f(v) w \, dx + \hat{\epsilon}_1(v) \hat{\epsilon}_{\text{vms}}^{\text{coer}}(v, w) \left(\nabla v, \nabla w\right)_{0,T} \\
&+ \int_{\partial T} [H(v) - f(v_T^+) \cdot n_T^+] w_T^+ \, ds \Big\}
\end{aligned} \tag{5.21}$$

folgt mit $H(v)-f(v_T^+)\cdot n_T^+ = \frac{1}{2}\left[f(v_T^-)-f(v_T^+)\right]\cdot n_T^+ + C_T(v_T^+-v_T^-)$ und $w=\eta'(v)\varphi$

$$\begin{aligned} & b(v,\eta'(v)\varphi) \\ &= \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\Big\{\int_T \nabla\cdot f(v)\eta'(v)\varphi\,dx + \hat{\epsilon}_1(v)\hat{\epsilon}_{\text{vms}}^{\text{coer}}(v,\eta'(v)\varphi)\left(\nabla v,\nabla(\eta'(v)\varphi)\right)_{0,T} \\ &\qquad + \int_{\partial T}\frac{1}{2}\left[f(v_T^-)-f(v_T^+)\right]\cdot n_T^+\eta'(v)\varphi\,ds + \int_{\partial T} C_T(v_T^+-v_T^-)\eta'(v)\varphi\,ds\Big\}. \end{aligned} \tag{5.22}$$

Hierbei bezeichnet $q=(\eta,q_1,\ldots,q_{d_x})$ das Entropie-Paar, wobei zusätzlich die Forderung

$$\eta'(0)=0 \tag{5.23}$$

erfüllt sein soll.
Mit der Kompatibilitätsbedingung (3.16) folgt

$$\nabla\cdot f(v)\eta'(v) = \sum_{i=0}^{d_x}\eta'(v)f_i'(v)v_{x_i} = \sum_{i=0}^{d_x} q_i'(v)v_{x_i} = \nabla\cdot q(v),$$

so dass sich

$$\int_T \nabla\cdot f(v)\eta'(v)\varphi\,dx = \int_T \nabla\cdot q(v)\varphi\,dx = -\int_T q(v)\cdot\nabla\varphi\,dx + \int_{\partial T} q(v)\cdot n_T\varphi\,ds$$

ergibt. Insgesamt resultiert

$$\begin{aligned} b(v,\eta'(v)\varphi) &= \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\Big\{\hat{\epsilon}_1(v)\hat{\epsilon}_{\text{vms}}^{\text{coer}}(v,\eta'(v)\varphi)\left(\nabla v,\nabla(\eta'(v)\varphi)\right)_{0,T} \\ &+ \int_{\partial T}\frac{1}{2}\left[f(v_T^-)-f(v_T^+)\right]\cdot n_T^+\eta'(v)\varphi\,ds + \int_{\partial T} C_T(v_T^+-v_T^-)\eta'(v)\varphi\,ds\Big\} \\ &+ \int_{\partial T} q(v)\cdot n_T\varphi\,ds - \int_{[0,t_N]\times\Omega} q(v)\cdot\nabla\varphi\,dx. \end{aligned} \tag{5.24}$$

Für die nächsten Überlegungen ist es wichtig, dass $\varphi_-^n(x)=\varphi_+^n(x)=\varphi^n(x)$ erfüllt ist. Mit $dx_x = dx_1dx_2\ldots dx_{d_x}$ folgt die Identität

$$\begin{aligned} &\sum_{T\in\mathcal{T}_h^n}\int_{\partial T} q(v)\cdot n_T\varphi\,ds \\ &= \sum_{\tau\in R_{n+1}}\int_\tau q(v)\cdot n\varphi\,ds + \sum_{\tau\in R_n}\int_\tau q(v)\cdot n\varphi\,ds \\ &+ \sum_{\tau\in R_{n,n+1}^i}\int_\tau \left[q(v^+)-q(v^-)\right]\cdot n^+\varphi\,ds + \sum_{\tau\in\Lambda_{n,n+1}}\int_\tau q(v^+)\cdot n^+\varphi\,ds \\ &= \int_\Omega \eta(v_-^{n+1})\varphi^{n+1}\,dx_x - \int_\Omega \eta(v_+^n)\varphi^n\,dx_x \\ &+ \sum_{\tau\in R_{n,n+1}^i}\int_\tau \left[q(v^+)-q(v^-)\right]\cdot n^+\varphi\,ds + \sum_{\tau\in\Lambda_{n,n+1}}\int_\tau q(v^+)\cdot n^+\varphi\,ds. \end{aligned}$$

Aufsummierung über n liefert

$$\begin{aligned}
&\sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \int_{\partial T} q(v)\cdot n_T \varphi \, ds \\
&= \int_\Omega \eta(v_-^N)\varphi^N \, dx_x - \int_\Omega \eta(v_-^0)\varphi^0 \, dx_x + \sum_{n=0}^{N-1} \int_\Omega \left[\eta(v_-^n) - \eta(v_+^n)\right] \varphi^n \, dx_x \\
&+ \sum_{n=0}^{N-1} \left\{ \sum_{\tau\in R_{n,n+1}^i} \int_\tau \left[q(v^+) - q(v^-)\right]\cdot n^+ \varphi \, ds + \sum_{\tau\in\Lambda_{n,n+1}} \int_\tau q(v^+)\cdot n^+ \varphi \, ds \right\}.
\end{aligned}$$

Schließlich folgt

$$\begin{aligned}
b(v,\eta'(v)\varphi) = &\sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \hat{\epsilon}_1(v)\hat{\epsilon}_{\text{vms}}^{\text{coer}}(v,\eta'(v)\varphi)\left(\nabla v, \nabla(\eta'(v)\varphi)\right)_{0,T} \\
&+ \int_\Omega \eta(v_-^N)\varphi^N \, dx_x - \int_\Omega \eta(v_-^0)\varphi^0 \, dx_x \\
&- \int_{[0,t_N]\times\Omega} q(v)\cdot\nabla\varphi \, dx \\
&+ \sum_{n=0}^{N-1} \int_\Omega \left[\eta(v_-^n) - \eta(v_+^n)\right] \varphi^n \, dx_x \\
&+ \sum_{n=0}^{N-1} \sum_{\tau\in R_{n,n+1}^i} \int_\tau \left[q(v^+) - q(v^-)\right]\cdot n^+ \varphi \, ds \\
&+ \sum_{n=0}^{N-1} \sum_{\tau\in\Lambda_{n,n+1}} \int_\tau q(v^+)\cdot n^+ \varphi \, ds \\
&+ \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \left\{ \int_{\partial T} \frac{1}{2}\left[f(v_T^-) - f(v_T^+)\right]\cdot n_T^+ \eta'(v)\varphi \, ds \right. \\
&\left. + \int_{\partial T} C_T(v_T^+ - v_T^-)\eta'(v)\varphi \, ds \right\}.
\end{aligned} \tag{5.25}$$

Damit können die beiden letzten Randintegrale analog behandelt werden:

$$
\begin{aligned}
&\sum_{T\in\mathcal{T}_h^n}\int_{\partial T}\frac{1}{2}\left[f(U_T^-)-f(U_T^+)\right]\cdot n_T^+\eta'(U)\varphi\,ds\\
&=\sum_{\tau\in R_{n+1}}\int_\tau\frac{1}{2}(U^- - U^+)\eta'(U)\varphi\,ds-\sum_{\tau\in R_n}\int_\tau\frac{1}{2}(U^- - U^+)\eta'(U)\varphi\,ds\\
&+\sum_{\tau\in R_{n,n+1}^i}\Bigg\{\int_\tau\frac{1}{2}\left[f(U^-)-f(U^+)\right]\cdot n^+\eta'(U^+)\varphi\,ds\\
&+\int_\tau\frac{1}{2}[f(U^+)-f(U^-)]\cdot n^-\eta'(U^-)\varphi\,ds\Bigg\}\\
&+\sum_{\tau\in\Lambda_{n,n+1}}\int_\tau\frac{1}{2}\left[f(g_D)-f(U^+)\right]\cdot n^+\eta'(U^+)\varphi\,ds\\
&=\int_\Omega\frac{1}{2}(U_+^{n+1}-U_-^{n+1})\eta'(U_-^{n+1})\varphi^{n+1}\,dx_x-\int_\Omega\frac{1}{2}(U_-^n-U_+^n)\eta'(U_+^n)\varphi^n\,dx_x\\
&+\sum_{\tau\in R_{n,n+1}^i}\int_\tau\frac{1}{2}\left[f(U^-)-f(U^+)\right]\cdot n^+\left[\eta'(U^+)+\eta'(U^-)\right]\varphi\,ds\\
&+\sum_{\tau\in\Lambda_{n,n+1}}\int_\tau\frac{1}{2}\left[f(g_D)-f(U^+)\right]\cdot n^+\eta'(U^+)\varphi\,ds,
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{T\in\mathcal{T}_h^n}\int_{\partial T}C_T(U_T^+-U_T^-)\eta'(U)\varphi\,ds\\
&=\sum_{\tau\in R_{n+1}}\int_\tau\frac{1}{2}(U^+-U^-)\eta'(U)\varphi\,ds+\sum_{\tau\in R_n}\int_\tau\frac{1}{2}(U^+-U^-)\eta'(U)\varphi\,ds\\
&+\sum_{\tau\in R_{n,n+1}^i}\left\{\int_\tau C_0^\Omega(U^+-U^-)\eta'(U^+)\varphi\,ds+\int_\tau C_0^\Omega(U^--U^+)\eta'(U^-)\varphi\,ds\right\}\\
&+\sum_{\tau\in\Lambda_{n,n+1}}\int_\tau C_0^{\partial\Omega}(U^+-g_D)\eta'(U^+)\varphi\,ds\\
&=\int_\Omega\frac{1}{2}(U_-^{n+1}-U_+^{n+1})\eta'(U_-^{n+1})\varphi^{n+1}\,dx_x+\int_\Omega\frac{1}{2}(U_+^n-U_-^n)\eta'(U_+^n)\varphi^n\,dx_x\\
&+\sum_{\tau\in R_{n,n+1}^i}\int_\tau C_0^\Omega(U^+-U^-)\left[\eta'(U^+)-\eta'(U^-)\right]\varphi\,ds\\
&+\sum_{\tau\in\Lambda_{n,n+1}}\int_\tau C_0^{\partial\Omega}(U^+-g_D)\eta'(U^+)\varphi\,ds.
\end{aligned}
$$

Letztendlich ergibt sich die Identität

$$\begin{aligned}
&\sum_{T\in\mathcal{T}_h^n}\left\{\int_{\partial T}\frac{1}{2}\left[f(U_T^-)-f(U_T^+)\right]\cdot n_T^+\eta'(U)\varphi\,ds+\int_{\partial T}C_T(U_T^+-U_T^-)\eta'(U)\varphi\,ds\right\}\\
&=-\int_\Omega(U_-^n-U_+^n)\eta'(U_+^n)\varphi^n\,dx_x\\
&+\sum_{\tau\in R_{n,n+1}^i}\int_\tau\frac{1}{2}\left[f(U^-)-f(U^+)\right]\cdot n^+\left[\eta'(U^+)+\eta'(U^-)\right]\varphi\,ds\\
&+\sum_{\tau\in R_{n,n+1}^i}\int_\tau C_0^\Omega(U^+-U^-)\left[\eta'(U^+)-\eta'(U^-)\right]\varphi\,ds\\
&+\sum_{\tau\in\Lambda_{n,n+1}}\int_\tau\frac{1}{2}\left[f(g_D)-f(0)-f(U^+)+f(0)\right]\cdot n^+\eta'(U^+)\varphi\,ds\\
&+\sum_{\tau\in\Lambda_{n,n+1}}\int_\tau C_0^{\partial\Omega}(U^+-g_D)\eta'(U^+)\varphi\,ds.
\end{aligned}\tag{5.26}$$

Setze

$$\begin{aligned}
E_0(f,\eta,v,\varphi)&=\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\hat{\epsilon}_1(v)\hat{\epsilon}_{\text{vms}}^{\text{coer}}(v,\eta'(v)\varphi)\left(\nabla v,\nabla(\eta'(v)\varphi)\right)_{0,T},\\
E_1(f,\eta,v,\varphi)&=\sum_{n=0}^{N-1}\int_\Omega\left[\eta(v_-^n)-\eta(v_+^n)-\eta'(v_+^n)(v_-^n-v_+^n)\right]\varphi^n\,dx_x
\end{aligned}\tag{5.27}$$

und

$$\begin{aligned}
E_2(f,\eta,v,\varphi)=\sum_{n=0}^{N-1}\sum_{\tau\in R_{n,n+1}^i}\int_\tau\Big[&q(v^+)-q(v^-)\\
&-\frac{1}{2}\left[f(v^+)-f(v^-)\right]\left[\eta'(v^+)+\eta'(v^-)\right]\Big]\cdot n^+\varphi\,ds.
\end{aligned}\tag{5.28}$$

Aus

$$\eta'(v^+)-\eta'(v^-)=(v^+-v^-)\int_0^1\eta''(v^-+r(v^+-v^-))dr$$

folgt

$$\begin{aligned}
E_3(f,\eta,v,\varphi)&=\sum_{n=0}^{N-1}\sum_{\tau\in R_{n,n+1}^i}\int_\tau C_0^\Omega(v^+-v^-)\left[\eta'(v^+)-\eta'(v^-)\right]\varphi\,ds\\
&=\sum_{n=0}^{N-1}\sum_{\tau\in R_{n,n+1}^i}\int_\tau C_0^\Omega(v^+-v^-)^2\left[\int_0^1\eta''(v^-+r(v^+-v^-))dr\right]\varphi\,ds.
\end{aligned}\tag{5.29}$$

Die Berücksichtigung der Kanten in $\Lambda_{n,n+1}$ erfolgt für

$$q(v)=\int_0^v\eta'(r)f'(r)\,dr$$

mit

$$E_4(f,\eta,v,\varphi) = \sum_{n=0}^{N-1} \sum_{\tau\in\Lambda_{n,n+1}} \int_\tau \left[q(v^+) - q(0) - \frac{1}{2}\left[f(v^+) - f(0)\right]\eta'(v^+)\right]\cdot n^+\varphi\, ds,$$

$$E_5(f,\eta,v,\varphi) = \sum_{n=0}^{N-1} \sum_{\tau\in\Lambda_{n,n+1}} \int_\tau C_0^{\partial\Omega} v^+\eta'(v^+)\varphi\, ds, \tag{5.30}$$

$$F_4(f,\eta,v,\varphi) = -\sum_{n=0}^{N-1} \sum_{\tau\in\Lambda_{n,n+1}} \int_\tau \frac{1}{2} f(g_D)\cdot n^+\eta'(v^+)\varphi\, ds \tag{5.31}$$

$$\tag{5.32}$$

und

$$F_5(f,\eta,v,\varphi) = \sum_{n=0}^{N-1} \sum_{\tau\in\Lambda_{n,n+1}} \int_\tau C_0^{\partial\Omega} g_D\eta'(v^+)\varphi\, ds. \tag{5.33}$$

Mit den neu eingeführten Bezeichnungen lässt sich die diskrete, stabilisierte Aufgabe (5.15) für $v = \eta'(U)\varphi$ schreiben als

$$\begin{aligned} b_1(U,\eta'(U)\varphi) + \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n} \left(\hat\epsilon_2(U) + \hat\epsilon_3(U)\right)\hat\epsilon_{\text{vms}}^{\text{coer}}(U,\eta'(U)\varphi)\left(\nabla U, \nabla(\eta'(U)\varphi)\right)_{0,T} \\ = \sum_{i=4}^{5} F_i(f,\eta,U,\varphi), \end{aligned} \tag{5.34}$$

wobei

$$\begin{aligned} b_1(v,\eta'(v)\varphi) = &\int_\Omega \eta(v_-^N)\varphi^N\, dx_x - \int_\Omega \eta(v_-^0)\varphi^0\, dx_x \\ &- \int_{[0,t_N]\times\Omega} q(v)\cdot\nabla\varphi\, dx \\ &+ \sum_{i=0}^{5} E_i(f,\eta,v,\varphi) \end{aligned} \tag{5.35}$$

ist.

Zusätzlich wird $a_1(\cdot,\cdot)$ ohne den entsprechenden Anteil der künstlichen Diffusion von $b_1(\cdot,\cdot)$ definiert:

$$\begin{aligned} a_1(v,\eta'(v)\varphi) = &\int_\Omega \eta(v_-^N)\varphi^N\, dx_x - \int_\Omega \eta(v_-^0)\varphi^0\, dx_x \\ &- \int_{[0,t_N]\times\Omega} q(v)\cdot\nabla\varphi\, dx \\ &+ \sum_{i=1}^{5} E_i(f,\eta,v,\varphi). \end{aligned} \tag{5.36}$$

Für die nachfolgenden Zwecke ist es wichtig, dass $\sum_{i=1}^{5} E_i(f,\eta,v,\varphi)$ nicht negativ ist. Die nächsten Überlegungen gewährleisten diese Forderung. Die Nichtnegativität von E_1 folgt direkt aus einer Eigenschaft über konvexe Funktionen [Alt02, Satz 3.2.1].
Betrachte jetzt

$$\begin{aligned}&\left[q(v^+) - q(v^-) - \frac{1}{2}\left[f(v^+) - f(v^-)\right]\left[\eta'(v^+) + \eta'(v^-)\right]\right]\cdot n^+\\ =&\left[\int_{v^-}^{v^+}\left[q' - f'\eta'\right]dr + \int_{v^-}^{v^+} f'\left(\eta' - \frac{1}{2}\left[\eta'(v^+) + \eta'(v^-)\right]\right)dr\right]\cdot n^+.\end{aligned}$$

Das erste Integral verschwindet wegen der Kompatibilitätsbedingung (3.16). Mit $[v] = v^+ - v^-$ und der Substitution $r = v^- + s[v]$ folgt

$$\begin{aligned}&\left[q(v^+) - q(v^-) - \frac{1}{2}\left[f(v^+) - f(v^-)\right]\left[\eta'(v^+) + \eta'(v^-)\right]\right]\cdot n^+\\ =&\int_0^1 f'(v^- + s[v])\cdot n^+\left(\eta'(v^- + s[v]) - \frac{1}{2}\left[\eta'(v^+) + \eta'(v^-)\right]\right)[v]\,ds.\end{aligned}$$

Die Konvexität von η bzw. die wachsende Monotonie von η' liefert

$$[\eta'(v^- + s[v]) - \eta'(v^-)][v] \geq 0$$

und

$$[\eta'(v^- + s[v]) - \eta'(v^+)][v] = [\eta'(v^+ - (1-s)[v]) - \eta'(v^+)][v] \leq 0.$$

Damit gilt

$$\begin{aligned}&\left|\left(\eta'(v^- + s[v]) - \frac{1}{2}\left[\eta'(v^+) + \eta'(v^-)\right]\right)[v]\right|\\ =&\left|\left(\frac{1}{2}[\eta'(v^- + s[v]) - \eta'(v^-)] + \frac{1}{2}[\eta'(v^- + s[v]) - \eta'(v^+)]\right)[v]\right|\\ \leq&\frac{1}{2}\left([\eta'(v^- + s[v]) - \eta'(v^-)] - [\eta'(v^+ - (1-s)[v]) - \eta'(v^+)]\right)[v] = \frac{1}{2}\mathrm{I}\end{aligned}$$

und schließlich

$$\begin{aligned}&\left|\left[q(v^+) - q(v^-) - \frac{1}{2}\left[f(v^+) - f(v^-)\right]\left[\eta'(v^+) + \eta'(v^-)\right]\right]\cdot n^+\right|\\ \leq&\frac{1}{2}\|f'\|_{0,\infty,\mathbb{R}}\int_0^1 \mathrm{I}\,ds.\end{aligned}$$

Die untere Schranke ergibt mit der Substitution $\xi = 1-s$ und durch (5.8)

$$\left[q(v^+) - q(v^-) - \frac{1}{2}\left[f(v^+) - f(v^-)\right]\left[\eta'(v^+) + \eta'(v^-)\right]\right]\cdot n^+ \geq$$

$$-\frac{1}{2}\|f'\|_{0,\infty,\mathbb{R}}\left(\int_0^1[\eta'(v^-+s[v])-\eta'(v^-)][v]\,ds-\int_0^1[\eta'(v^+-\xi[v])-\eta'(v^+)][v]\,d\xi\right)$$

sowie nach Anwendung des Mittelwertsatzes

$$\begin{aligned}
&\left[q(v^+)-q(v^-)-\frac{1}{2}\left[f(v^+)-f(v^-)\right]\left[\eta'(v^+)+\eta'(v^-)\right]\right]\cdot n^+\\
&\geq-\frac{1}{2}\|f'\|_{0,\infty,\mathbb{R}}\int_0^1\int_0^1[\eta''(v^-+ts[v])+\eta''(v^+-ts[v])][v]^2sdtds\\
&=-\frac{1}{2}[v]^2\|f'\|_{0,\infty,\mathbb{R}}\int_0^1\int_0^s[\eta''(v^-+r[v])+\eta''(v^+-r[v])]\,drds\\
&\geq-\frac{1}{2}[v]^2\|f'\|_{0,\infty,\mathbb{R}}\int_0^1\int_0^1[\eta''(v^-+r[v])+\eta''(v^+-r[v])]\,drds\\
&=-\frac{1}{2}[v]^2\|f'\|_{0,\infty,\mathbb{R}}\int_0^1[\eta''(v^-+r[v])+\eta''(v^+-r[v])]\,dr\\
&=-[v]^2\|f'\|_{0,\infty,\mathbb{R}}\int_0^1\eta''(v^-+r[v])\,dr,
\end{aligned}$$

wobei $\eta''\geq 0$ und $\int_0^1\eta''(v^+-r[v])dr=\int_0^1\eta''(v^-+r[v])dr$ benutzt wurde. E_2 lässt sich somit für nichtnegative φ wie folgt abschätzen:

$$E_2(f,\eta,v,\varphi)\geq-\sum_{n=0}^{N-1}\sum_{\tau\in R^i_{n,n+1}}\int_\tau\|f'\|_{0,\infty,\mathbb{R}}\left(\int_0^1\eta''(v^-+r[v])\,dr\right)[v]^2\varphi\,ds.$$

Zusammen mit der Definition (5.29) für E_3 ergibt sich nun

$$\begin{aligned}
&E_2(f,\eta,v,\varphi)+E_3(f,\eta,v,\varphi)\\
&\geq\sum_{n=0}^{N-1}\sum_{\tau\in R^i_{n,n+1}}\int_\tau\left(C_0^\Omega-\|f'\|_{0,\infty,\mathbb{R}}\right)\left(\int_0^1\eta''(v^-+r[v])\,dr\right)[v]^2\varphi\,ds,
\end{aligned}\tag{5.37}$$

so dass mit der Annahme (5.8) jetzt

$$E_2(f,\eta,v,\varphi)+E_3(f,\eta,v,\varphi)\geq 0$$

gewährleistet ist.
Wegen (5.23) können E_4 und E_5 analog zu E_2 und E_3 behandelt werden:

$$\begin{aligned}
&E_4(f,\eta,v,\varphi)+E_5(f,\eta,v,\varphi)\\
&\geq\sum_{n=0}^{N-1}\sum_{\tau\in\Lambda_{n,n+1}}\int_\tau\left(C_0^{\partial\Omega}-\|f'\|_{0,\infty,\cdot|_{\partial\Omega}}\right)\int_0^1\eta''(rv^+)\,dr(v^+)^2\varphi\,ds\\
&\geq\sum_{n=0}^{N-1}\sum_{\tau\in\Lambda_{n,n+1}}\int_\tau\frac{1}{2}C_0^{\partial\Omega}\int_0^1\eta''(rv^+)\,dr(v^+)^2\varphi\,ds\geq 0.
\end{aligned}\tag{5.38}$$

Betrachte jetzt ein weiteres Mal die diskrete, stabilisierte Aufgabe (5.15) bzw. (5.34) speziell mit $\eta(v) = v^2/2$, $v = U$, $\varphi = 1$:

$$b_1(U,U) + \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \left(\hat{\epsilon}_2(U) + \hat{\epsilon}_3(U)\right) \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U,U) \left(\nabla U, \nabla U\right)_{0,T} = \sum_{i=4}^{5} F_i(f, U^2/2, U, 1). \tag{5.39}$$

Insgesamt gilt damit die Identität

$$\begin{aligned} b_1(U,U) &= b_1(U, \eta'(U)\varphi) \\ &= \int_\Omega \eta(U_-^N)\varphi^N \, dx_x - \int_\Omega \eta(U_-^0)\varphi^0 \, dx_x - \int_{[0,t_N]\times\Omega} q(U) \cdot \nabla\varphi \, dx \\ &\quad + \sum_{i=0}^{5} E_i(f, \eta, U, \varphi) \\ &= \frac{1}{2}\int_\Omega (U_-^N)^2 \, dx_x - \frac{1}{2}\int_\Omega (U_-^0)^2 \, dx_x + \sum_{i=0}^{5} E_i(f, U^2/2, U, 1). \end{aligned} \tag{5.40}$$

Aufgrund der Nichtnegativität der Terme $\sum_{i=1}^{3} E_i(f, U^2/2, U, 1)$ erlaubt (5.39), (5.40) und $U_-^0 = u_0$ die Aufstellung der Ungleichung

$$\begin{aligned} \frac{1}{2}\int_\Omega (U_-^N)^2 \, dx_x + \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_1(U) \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U,U) \left(\nabla U, \nabla U\right)_{0,T} + \frac{1}{2} C_0^{\partial\Omega} \|U^+\|_{0,2,\Sigma_T}^2 \\ \leq \frac{1}{2}\int_\Omega (u_0)^2 \, dx_x + \sum_{i=4}^{5} F_i(f, U^2/2, U, 1). \end{aligned} \tag{5.41}$$

Abschließend folgt die Behandlung der Randterme $F_i(f, U^2/2, U, 1)$, $i = 4, 5$.

$$\begin{aligned} & F_4(f, U^2/2, U, 1) + F_5(f, U^2/2, U, 1) \\ = & \sum_{n=0}^{N-1} \sum_{\tau \in \Lambda_{n,n+1}} - \int_\tau \frac{1}{2} f(g_D) \cdot n^+ U^+ \, ds + \int_\tau C_0^{\partial\Omega} g_D U^+ \, ds \\ \underset{(5.9)}{\leq} & \sum_{n=0}^{N-1} \sum_{\tau \in \Lambda_{n,n+1}} \int_\tau \left(\int_0^1 \frac{1}{2} \left| f'(r g_D) \cdot n^+ \right| dr + C_0^{\partial\Omega} \right) \left| g_D U^+ \right| \, ds \\ \leq & \int_{\Sigma_T} \frac{1}{2} \left(\int_0^1 \frac{1}{2} \left| f'(r g_D) \cdot n^+ \right| dr + C_0^{\partial\Omega} \right)^2 \frac{1}{C_0^{\partial\Omega}} g_D^2 \, ds + \int_{\Sigma_T} \frac{1}{2} C_0^{\partial\Omega} (U^+)^2 \, ds \\ \leq & \frac{25}{32} C_0^{\partial\Omega} \|g_D\|_{0,2,\Sigma_T}^2 + \frac{1}{2} C_0^{\partial\Omega} \|U^+\|_{0,2,\Sigma_T}^2. \end{aligned} \tag{5.42}$$

Es folgt

$$\begin{aligned} \frac{1}{2}\|U_-^N\|_{0,2,\Omega}^2 + \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}(U) \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U,U) \|\nabla U\|_{0,2,T}^2 + \frac{1}{2} C_0^{\partial\Omega} \|U^+\|_{0,2,\Sigma_T}^2 \\ \leq \frac{1}{2}\|u_0\|_{0,2,\Omega}^2 + C_0^{\partial\Omega} \|g_D\|_{0,2,\Sigma_T}^2 + \frac{1}{2} C_0^{\partial\Omega} \|U^+\|_{0,2,\Sigma_T}^2, \end{aligned} \tag{5.43}$$

bzw.

$$\begin{aligned}\frac{1}{2}\|U_-^N\|_{0,2,\Omega}^2 + \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\hat{\epsilon}(U)\hat{\epsilon}_{\text{vms}}^{\text{coer}}(U,U)\|\nabla U\|_{0,2,T}^2 \\ \leq \frac{1}{2}\|u_0\|_{0,2,\Omega}^2 + C_0^{\partial\Omega}\|g_D\|_{0,2,\Sigma_T}^2 = C_5.\end{aligned} \tag{5.44}$$

Das nächste Ziel ist, eine Abschätzung der diskreten Lösung für alle $t \in (0,T)$ zu erhalten. Betrachte dazu die Darstellung von einem festem $X \in \Omega$ als Startpunkt einer Charakteristik $x = x(t,t_s,X)$ zum Zeitpunkt $t = t_s$. Es wird angenommen, dass x lokal existiert und eindeutig ist. Für (3.1) ist $x(t,t_s,X) = f_x{}'(u(t_s,X))(t-t_s)+X$. Somit gilt für ein fixes t', $p \in \mathbb{N}$ und $F(X) = x(t',t_s,X)$

$$\begin{aligned}\|v(t',\cdot)\|_{0,p,\Omega}^p &= \int_\Omega v(t',y)^p\,dy = \left|\begin{smallmatrix} y &= & F(X) \\ dy &= & |\partial F/\partial X|dX=dX\end{smallmatrix}\right| \\ &= \int_{F^{-1}(\Omega)} v(t',x(t',t_s,X))^p\,dX.\end{aligned} \tag{5.45}$$

Für $t_{N-1} \leq t \leq t_N$ gilt damit

$$\begin{aligned}\|U(t,\cdot)\|_{0,p,\Omega}^p &= \|U_-^N\|_{0,p,\Omega}^p - \int_t^{t_N}\frac{\partial}{\partial t'}\|U(t',\cdot)\|_{0,p,\Omega}^p\,dt' \\ &= \|U_-^N\|_{0,p,\Omega}^p - \int_t^{t_N}\int_{F^{-1}(\Omega)}\frac{\partial}{\partial t'}U(t',x(t',t_s,X))^p\,dX\,dt' \\ &= \|U_-^N\|_{0,p,\Omega}^p - p\int_t^{t_N}\int_{F^{-1}(\Omega)}U^{p-1}(U_t+\nabla U\cdot\dot{x})\,dX\,dt' \\ &= \|U_-^N\|_{0,p,\Omega}^p - p\int_t^{t_N}\int_{F^{-1}(\Omega)}U^{p-1}(U_t+\nabla U\cdot f_x{}'(U))\,dX\,dt' \\ = \left|\begin{smallmatrix} y &= F(X) \\ dy &= dX\end{smallmatrix}\right| &= \|U_-^N\|_{0,p,\Omega}^p - p\int_t^{t_N}\int_\Omega U^{p-1}(U_t+\nabla U\cdot f_x{}'(U)\,dy\,dt' \\ = |\,{}^{x_x\,=\,y}\,| &= \|U_-^N\|_{0,p,\Omega}^p - p\int_t^{t_N}\int_\Omega U^{p-1}(t',x_x)\ \text{div}\ f(U(t',x_x))\,dx_x\,dt'.\end{aligned} \tag{5.46}$$

Vermöge dieser Darstellung erfolgt für $p = 2$ die Abschätzung der diskreten Lösung in der

L^2-Norm:

$$\begin{aligned}
\|U(t,\cdot)\|_{0,2,\Omega}^2 &\leq \|U_-^N\|_{0,2,\Omega}^2 + 2\int_t^{t_N}\int_\Omega |U(t',x_x)\ \mathrm{div}\ f(U(t',x_x))|\ dx_x dt' \\
&\leq \|U_-^N\|_{0,2,\Omega}^2 + 2\int_t^{t_N}\int_\Omega \frac{C_1 \min\limits_{T\in\mathcal{T}_h^{N-1}}\{h_T\}}{\|f'(U)\|_{l^2}} |\operatorname{div} f(U)|^2\, dx_x dt' \\
&\quad + 2\int_t^{t_N}\int_\Omega \frac{\|f'(U)\|_{l^2}}{4C_1 \min\limits_{T\in\mathcal{T}_h^{N-1}}\{h_T\}} U^2\, dx_x dt' \\
&\leq \|U_-^N\|_{0,2,\Omega}^2 + 2\int_t^{t_N}\int_\Omega \frac{C_1 \min\limits_{T\in\mathcal{T}_h^{N-1}}\{h_T\}}{\|f'(U)\|_{l^2}} \|f'(U)\|_{l^2}^2 \|\nabla U\|_{l^2}^2\, dx_x dt' \\
&\quad + \frac{\|f'\|_{0,\infty,\mathbb{R}}}{2C_1 \min\limits_{T\in\mathcal{T}_h^{N-1}}\{h_T\}} \int_t^{t_N} \|U(t',\cdot)\|_{0,2,\Omega}^2\, dt' \\
&\leq \|U_-^N\|_{0,2,\Omega}^2 + 2C_1 \sum_{T\in\mathcal{T}_h^{N-1}} h_T \max_{x\in T}\|f'(U)\|_{l^2} \int_T \nabla U\cdot\nabla U\, dx \\
&\quad + \frac{C_0^\Omega}{2C_1 \min\limits_{T\in\mathcal{T}_h^{N-1}}\{h_T\}} \int_t^{t_N} \|U(t',\cdot)\|_{0,2,\Omega}^2\, dt'.
\end{aligned} \tag{5.47}$$

Insgesamt gilt mit (5.8) und der Gronwallschen Ungleichung (Lemma 2.3.11)

$$\begin{aligned}
&\|U(t,\cdot)\|_{0,2,\Omega}^2 \\
&\leq \|U_-^N\|_{0,2,\Omega}^2 + \frac{C_0^\Omega}{2C_1 \min\limits_{T\in\mathcal{T}_h^{N-1}}\{h_T\}} \int_t^{t_N} \|U(t',\cdot)\|_{0,2,\Omega}^2\, dt' \\
&\quad + 2C_1 \sum_{T\in\mathcal{T}_h^{N-1}} h_T \max_{x\in T}\|f'(U)\|_{l^2} \int_T \nabla U\cdot\nabla U\, dx \\
&\leq C_6 \left[2C_1 \sum_{T\in\mathcal{T}_h^{N-1}} h_T \max_{x\in T}\|f'(U)\|_{l^2} \|\nabla U\|_{0,2,T}^2 + \|U_-^N\|_{0,2,\Omega}^2 \right] \\
&\leq \frac{C_6}{\hat{\epsilon}_{\mathrm{vms}}^{\min}} \left[2C_1 \sum_{T\in\mathcal{T}_h^{N-1}} h_T \max_{x\in T}\|f'(U)\|_{l^2} \hat{\epsilon}_{\mathrm{vms}}^{\mathrm{coer}}(U,U) \|\nabla U\|_{0,2,T}^2 + \|U_-^N\|_{0,2,\Omega}^2 \right]
\end{aligned} \tag{5.48}$$

mit $\exp(\frac{C_0^\Omega}{2C_1 \min\limits_{T\in\mathcal{T}_h^{N-1}}\{h_T\}}(t_N - t)) < \exp(\frac{C_0^\Omega}{C_1 C_{qu}}) = C_6$, falls $\mathcal{T}_h^{N-1}$ quasiuniform ist. (5.44) liefert nun das folgende Theorem:

Theorem 5.3.1 *Es sei Ω ein Gebiet mit Lipschitz-Rand und $\left\{\mathcal{T}_h^{N-1}\right\}_{h>0}$ eine quasiuniforme Familie von Partitionierungen von $(t_{N-1}, t_N)\times\Omega$. Dann existiert ein C unabhängig von h, so*

dass für die Lösung U von (5.15) zusammen mit der Annahme (5.8) und (5.9) die Ungleichung

$$\|U(t,\cdot)\|_{0,2,\Omega} \leq \left(\frac{C_6}{\hat{\epsilon}^{\min}_{\text{vms}}}\right)^{1/2} \left[\|u_0\|_{0,2,\Omega} + \sqrt{2C_0^{\partial\Omega}}\|g_D\|_{0,2,\Sigma_T}\right] \leq C \tag{5.49}$$

für $t \in (t_{N-1}, t_N)$ gilt.

5.4 $L^\infty(L^\infty)$-Abschätzung der diskreten Lösung

5.4.1 Koerzivität des Shock-capturing Terms für v und $I_h^k(v^{p-1})$

Die Schwierigkeit für $p = 2m,\ m \in \mathbb{N}$ eine Ungleichung der Form

$$\left(\nabla v, \nabla I_h^k(v^{p-1})\right)_{0,T} \geq C \int_T \|\nabla v\|_{l^2}^2 \|v\|_{0,\infty,T}^{p-2}\, dx$$

mit einem $C > 0$ unabhängig von h zu gewährleisten, liegt für $p \neq 2$ in der Verschiedenheit der Argumente. Vergleichbare Ergebnisse sind dem Autor bisher nur aus [Sze89a, Lemma 4.2] und [Sze91, Lemma 3.3] bekannt. Allerdings gelten diese Aussagen nur für spezielle Dreiecke bzw. Tetraeder und auch nur für lineare Ansatzfunktionen. Darüber hinaus enthält die Koerzivitätskonstante C in diesen Abschätzungen noch den Faktor p. Die in diesem Abschnitt gewählte Vorgehensweise beseitigt alle genannten Defizite. Vorausgesetzt wird lediglich, dass die lokale Diskretisierungsmatrix A des Shock-capturing Terms den einfachen Eigenwert $\lambda = 0$ zum Eigenvektor $(1, \cdots, 1)^T$ besitzt und überdies symmetrisch, positiv semidefinit ist.

Für diesen Zweck wird zunächst die benötigte Notation und Theorie zum Thema *numerischer Wertebereich* für lineare Operatoren in *Banach-Räumen* bereitgestellt.

Definition 5.4.1 Sei $(X, \|\cdot\|)$ ein normierter Raum, $S(X)$ die Einheitssphäre und X' der *Dualraum* von X. Für einen linearen Operator A auf X, ist

$$W(A,\ \|\cdot\|) = \{f(Ax) :\ (x,f) \in \Pi\} \tag{5.50}$$

der *räumliche numerische Wertebereich* mit $\Pi = \{(x,f) \in S(X) \times S(X') :\ f(x) = 1\}$.

Die Existenz eines $f \in X'$ für ein $x \in X$ mit $\|x\| = 1$ ist durch eine Folgerung aus dem *Fortsetzungssatz von Hahn-Banach* (vgl. [Heu92, Satz 36.4]) gewährleistet. Im Gegensatz zum *Spektrum* $\sigma(A)$ hängt der Wertebereich $W(A,\ \|\cdot\|)$ auch von der verwendeten Norm ab.

Bemerkung 5.4.2 Diese Definition des Wertebereichs ist äquivalent zu

$$W(A,\ \|\cdot\|) = \{f(Ax) :\ f(x) = \|x\|\|f\| = 1\}, \tag{5.51}$$

denn es gilt:

$$\|f\| = \sup_{\|x\|=1} |f(x)| \leq \|f\| \underbrace{\|x\|}_{=1} = 1.$$

Der für die Koerzivitätsabschätzung geeignete Raum $(\mathbb{R}^n, \|\cdot\|_{l^p})$ liefert dank der Darstellung $f(x) = \sum_{i=1}^n x_i f(e_i) = x^T y_f$ und der Normisomorphie von $(l^p)'$ und l^q für $1/p + 1/q = 1$ die Identität $\|f\| = \|y_f\|_{l^q}$.

Die Bedingung in der äquivalenten Definition des Wertebereichs ist somit die Identität der Hölderschen Ungleichung $x^T y_f = \|x\|_{l^p}\|y_f\|_{l^q} = 1$, in Zeichen $x \parallel y_f$ (vgl. [Bau62]).

Für den Spezialfall $p = q = 2$ entsteht der bekannte numerische Wertebereich $W(A)$ nach Toeplitz [Toe18] für lineare Operatoren auf *Hilbert-Räumen*. $W(A, \|\cdot\|)$ ist im Gegensatz zu $W(A)$ nicht notwenig konvex ([NS64, S. 357]).

Ein zweiter numerischer Bereich kann definiert werden, indem die Matrix A als Element einer *normierten Algebra* $\mathcal{A}$ mit Einselement aufgefasst wird.

Definition 5.4.3 Sei $\mathcal{A}$ eine normierte Algebra, $S(\mathcal{A}) = \{x \in \mathcal{A} : \|x\| = 1\}$ die Einheitssphäre und $\mathcal{A}'$ ist der Dualraum von $\mathcal{A}$. Für ein $x \in \mathcal{A}$ ist

$$D(\mathcal{A}, x) = \{f \in \mathcal{A}' : f(x) = 1 = \|f\|\} \tag{5.52}$$

und

$$V_{\mathcal{A}}(a, x, \|\cdot\|) = \{f(ax) : f \in D(\mathcal{A}, x)\}. \tag{5.53}$$

Der *algebraische numerische Wertebereich* wird definiert als (vgl. [BD71, S. 15])

$$V_{\mathcal{A}}(a, \|\cdot\|) = \cup\{V_{\mathcal{A}}(a, x, \|\cdot\|) : x \in S(\mathcal{A})\}. \tag{5.54}$$

Für den algebraischen numerischen Wertebereich genügt es, nur das Einselement zu betrachten, denn es gilt

Lemma 5.4.4

$$V_{\mathcal{A}}(a, \|\cdot\|) = V_{\mathcal{A}}(a, 1, \|\cdot\|),\ a \in \mathcal{A}.$$

Beweis [BD71, Lemma 2.2]. □

Der Zusammenhang zwischen $V_{\mathcal{A}}$ und W ist mit dem folgenden Lemma gegeben, sofern die Matrix A als $a \in \mathcal{A}$ aufgefasst wird:

Lemma 5.4.5

$$\operatorname{conv} W(A, \|\cdot\|) = V_{\mathcal{A}}(a, \|\cdot\|). \tag{5.55}$$

Beweis [BD71, S. 84] oder [LS04, Corollary 2.2]. □

Für *hermitesche* ($V_{\mathcal{A}}(a, \|\cdot\|) \subset \mathbb{R}$) Elemente a der komplexen normierten Algebra $\mathcal{A}$ gilt ferner

Theorem 5.4.6 (Vidav) *Sei $a \in \mathcal{A}$ hermitesch. Dann gilt:*

$$\operatorname{conv} \sigma(a) = V_{\mathcal{A}}(a, \|\cdot\|). \tag{5.56}$$

Beweis [BD71, Corollary 5.11]. □

Korollar 5.4.7 *Sei A eine symmetrische, positiv semidefinite Matrix. Dann gilt:*

$$W(A, \|\cdot\|) \subseteq [0, \lambda_{\max}(A)], \tag{5.57}$$

wobei $\lambda_{\max}(A)$ den größten Eigenwert der Matrix A bezeichnet.

Mit dem letzten Korollar ist es nun möglich, die Koerzivität des Shock-capturing Terms zu gewährleisten:

Lemma 5.4.8 *Unter den Voraussssetzungen der Definition (4.50) und für Lagrangesche Finite-Elemente gilt für alle $v \in W_h$ und $p = 2m$, $m \in \mathbb{N}$:*

$$\left(\nabla v, \nabla I_h^k(v^{p-1})\right)_{0,T} \geq \frac{\lambda_{\min}(A)}{(k+1)^d \Lambda_k \lambda_{\max}(A)} \int_T \|\nabla v\|_{l^2}^2 \|v\|_{0,\infty,T}^{p-2}\, dx \tag{5.58}$$

und

$$\left(P_h'(\nabla v), P_h'(\nabla I_h^k(v^{p-1})\right)_{0,T,h} \geq 0, \tag{5.59}$$

mit $A_{ij} = (\nabla\varphi_j, \nabla\varphi_i)_{0,\hat{T}}$, $1 \leq i,j \leq n_{dof}^k$ und $\Lambda_k = \| \sum_{i=1}^{n_{dof}^k} |\varphi_i| \|_{0,\infty,\hat{T}}$ der Lebesgue-Konstanten. $\lambda_{\min}(A)$ ist der kleinste, positive Eigenwert von A und $\lambda_{\max}(A)$ der größte Eigenwert.

Beweis Der Beweis von (5.58) erfolgt auf dem Referenzelement $\hat{T}$ und besitzt dank eines Homogenitätsarguments und $\|v\|_{0,\infty,T} = \|\hat{v}\|_{0,\infty,\hat{T}}$ auch Gültigkeit auf T. (5.59) folgt analog.
Für $v = \text{const}$ ist die Ungleichung trivial erfüllt. Sei also zunächst $v \neq \text{const}$. Es erfolgt die Aufteilung von $\mathbb{Q}_k(\hat{T}) = V^0(\hat{T}) \oplus V(\hat{T})$ mit

$$\begin{aligned} V^0(\hat{T}) &= \{v \in \mathbb{Q}_k(\hat{T}) \setminus \{0\} : v = \text{const}\} \\ &= \{v \in \mathbb{Q}_k(\hat{T}) \setminus \{0\} : \int_{\hat{T}} \nabla w \cdot \nabla v\, dx = 0,\ w \in \mathbb{Q}_k(\hat{T})\}. \end{aligned}$$

Für Lagrangesche Finite-Elemente ist der Koeffizientenraum von V^0 somit

$$V_{\mathcal{N}}^0 = \text{span}\{(1,\ldots,1)^T\} \setminus \{0\},\ \dim V_{\mathcal{N}}^0 = 1. \tag{5.60}$$

Betrachte jetzt mit $\nabla\varphi = (\nabla\varphi_1, \ldots, \nabla\varphi_{n_{\text{dof}}^k})^T$, $v_{\mathcal{N}} = (v(x))_{x\in\mathcal{N}}$, $v_{\mathcal{N}}^{p-1} = (v^{p-1}(x))_{x\in\mathcal{N}}$:

$$\frac{\left(\nabla v, \nabla I_h^k(v^{p-1})\right)_{0,\hat{T}}}{\|v_{\mathcal{N}}\|_{l^p}^p} = \frac{v_{\mathcal{N}}^T A v_{\mathcal{N}}^{p-1}}{v_{\mathcal{N}}^T v_{\mathcal{N}}^{p-1}} = v_{\mathcal{N}}^T A v_{\mathcal{N}}^{p-1},$$

falls $v_{\mathcal{N}}^T v_{\mathcal{N}}^{p-1} = 1$. Diese Normierung ist dank der Homogenität des Quotienten jedoch stets möglich. Durch Nachrechnen folgt weiter $1 = v_{\mathcal{N}}^T v_{\mathcal{N}}^{p-1} = \|v_{\mathcal{N}}\|_{l^p} \|v_{\mathcal{N}}^{p-1}\|_{l^{p/(p-1)}}$ und somit

$$v_{\mathcal{N}}^T A v_{\mathcal{N}}^{p-1} \in W(A, \|\cdot\|_{l^p}) \underset{(5.57)}{\subseteq} [0, \lambda_{\max}(A)].$$

Die Matrix A ist symmetrisch, positiv semidefinit, wobei die Eigenwerte der Größe nach sortiert seien:

$$0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_{n_{\text{dof}}^k}.$$

Der Eigenraum zum Eigenwert 0 ist nach Definition der Matrix A gerade $V_{\mathcal{N}}^0$. Die Berücksichtigung der Tatsache $v_{\mathcal{N}} \neq \text{const}$ mit $v_{\mathcal{N}}^T v_{\mathcal{N}}^{p-1} = 1$ gelingt, da eine spektrale Zerlegung der Matrix A in dyadische Produkte der Eigenvektoren $\xi_i, 1 \leq i \leq n_{\text{dof}}^k$ möglich ist (vgl. z.B. [ZF84, S. 274]):

$$v_{\mathcal{N}}^T A v_{\mathcal{N}}^{p-1} = v_{\mathcal{N}}^T \sum_{i=1}^{n_{\text{dof}}^k} \lambda_i \xi_i \xi_i^T v_{\mathcal{N}}^{p-1} = \sum_{i=1}^{n_{\text{dof}}^k} \lambda_i v_{\mathcal{N}}^T \xi_i \xi_i^T v_{\mathcal{N}}^{p-1} = \sum_{i=2}^{n_{\text{dof}}^k} \lambda_i v_{\mathcal{N}}^T \xi_i \xi_i^T v_{\mathcal{N}}^{p-1}.$$

Zunächst gilt mit (5.57) : $v_{\mathcal{N}}^T \xi_i \xi_i^T v_{\mathcal{N}}^{p-1} \in [0,1]$. Da allerdings $v_{\mathcal{N}} \neq \text{const}$ ist, sind

$$v_{\mathcal{N}}, v_{\mathcal{N}}^{p-1} \notin \text{span}\{\xi_1\}$$

und somit ist $v_{\mathcal{N}}^T \xi_i \xi_i^T v_{\mathcal{N}}^{p-1} > 0$ für wenigstens ein $i = 2, \ldots, n_{\text{dof}}^k$. Insgesamt folgt

$$\frac{\left(\nabla v, \nabla I_h^k(v^{p-1})\right)_{0,\hat{T}}}{\|v_{\mathcal{N}}\|_{l^p}^p} \geq \lambda_2, \quad v_{\mathcal{N}} \neq \text{const}\,. \tag{5.61}$$

Ferner gilt

$$\begin{aligned}
\|v\|_{0,\infty,\hat{T}}^{p-2} \|\nabla v\|_{0,2,\hat{T}}^2 &\underset{(2.26)}{\leq} \lambda_{\max}(A) \|v^{p-2}\|_{0,\infty,\hat{T}} \|v_{\mathcal{N}}\|_{l^2}^2 \\
&\leq \lambda_{\max}(A) \Lambda_k \|v_{\mathcal{N}}\|_{l^\infty}^{p-2} \|v_{\mathcal{N}}\|_{l^2}^2 \\
&\leq \lambda_{\max}(A) \Lambda_k \|v_{\mathcal{N}}\|_{l^p}^{p-2} \|v_{\mathcal{N}}\|_{l^2}^2 \\
&\underset{(2.6)}{\leq} \left(n_{\text{dof}}^k\right)^{1-2/p} \lambda_{\max}(A) \Lambda_k \|v_{\mathcal{N}}\|_{l^p}^p \\
&\leq (k+1)^d \frac{\lambda_{\max}(A)}{\lambda_2} \Lambda_k \left(\nabla v, \nabla I_h^k(v^{p-1})\right)_{0,\hat{T}}\,. \quad \square
\end{aligned}$$

Lemma 5.4.9 *Es existiert eine Konstante $C > 0$ unabhängig von h, k und $p \in \mathbb{N}$ derart, dass gilt:*

$$|v^{p-1}|_{k+1,\infty,T} \leq C p^{k+1} \left(\frac{h}{k^2}\right)^{-k+1} \|\nabla v\|_{0,\infty,T}^2 \|v\|_{0,\infty,T}^{p-3} \tag{5.62}$$

für $p \geq 3$ und alle $v \in W_h$ mit $v_{Q_{n,n+1}}|_T \in P(T) \quad \forall T \in \mathcal{T}_h^n$.

Beweis Induktion über k : Für $k = 1$ ist die Ungleichung erfüllt, denn es gilt

$$\begin{aligned}
\partial_{ij} v^{p-1} &= \partial_i(\partial_j v^{p-1}) = (p-1)\partial_i(v^{p-2}\partial_j v) \\
&= (p-1)\left[(p-2)v^{p-3}\partial_i v \partial_j v + v^{p-2}\partial_{ij} v\right] \\
&= (p-1)(p-2)v^{p-3}\partial_i v \partial_j v.
\end{aligned}$$

Annahme: (5.62) ist wahr für ein $k \geq 1$. Sei β ein Multiindex mit $|\beta| = k+1$ und $0 \leq i \leq d_x$. Dann gilt mit der Leibnizregel (2.31):

$$\partial_i \partial^\beta (v^{p-1}) = (p-1)\partial^\beta \left(v^{p-2}\partial_i v\right) = (p-1) \sum_{\alpha \leq \beta} \binom{\beta}{\alpha} \partial^\alpha (v^{p-2}) \partial^{\beta-\alpha} \partial_i v.$$

(5.62) liefert

$$\|\partial^\alpha \left(v^{p-2}\right)\|_{0,\infty,T} \leq C(p-1)^{|\alpha|}(h/k^2)^{2-|\alpha|}\|\nabla v\|^2_{0,\infty,T}\|v\|^{p-4}_{0,\infty,T},$$

und mit der inversen Ungleichung (4.61) :

$$\|\partial^{\beta-\alpha}\partial_i v\|_{0,\infty,T} \leq C(h/k^2)^{-|\beta-\alpha|-1}\|v\|_{0,\infty,T}$$

folgt wegen $\alpha \leq \beta \Rightarrow |\beta-\alpha| = |\beta| - |\alpha|$

$$\begin{aligned}\|\partial_i\partial^\beta \left(v^{p-1}\right)\|_{0,\infty,T} &\leq C(p-1)\sum_{\alpha\leq\beta}\binom{\beta}{\alpha}(p-1)^{|\alpha|}\left(\frac{h}{k^2}\right)^{1-|\beta|}\|\nabla v\|^2_{0,\infty,T}\|v\|^{p-3}_{0,\infty,T}\\ &\leq C(p-1)^{k+2}\left(\frac{h}{k^2}\right)^{1-(k+1)}\|\nabla v\|^2_{0,\infty,T}\|v\|^{p-3}_{0,\infty,T}.\end{aligned}$$

Also gilt (5.62) auch für $\tilde{k} = k+1$. □

5.4.2 Diskretisierung des Problems mit der Testfunktion $I_h^k(v^{p-1})$

In diesem Abschnitt wird schließlich die gleichmäßige Beschränktheit bzgl. der Gitterweite h der diskreten Lösung von (5.15) in der $L^\infty(L^\infty)$-Norm bewiesen. Ausgangspunkt ist die Formulierung (5.34) mit $\eta(v) = \frac{1}{p}v^p$ und $\varphi = 1$. Da $\eta'(U)\varphi = U^{p-1}$ für $p \neq 2$ nicht in dem Finite-Elemente-Raum W_h liegt, wird die schwache Formulierung mit $v = I_h^k(U^{p-1})$ getestet:

$$b(U, I_h^k(U^{p-1})) + \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\left(\hat{\epsilon}_2(U)+\hat{\epsilon}_3(U)\right)\hat{\epsilon}^{\text{coer}}_{\text{vms}}(U, I_h^k(U^{p-1}))\left(\nabla U, \nabla I_h^k(U^{p-1})\right)_{0,T} = 0. \tag{5.63}$$

Die zentrale Idee des Beweises ist es den Interpolationsfehler $U^{p-1} - I_h^k(U^{p-1})$ mit Hilfe der Linearität im zweiten Argument von $b(\cdot,\cdot)$ durch die spezielle Bauart des Shock-capturing Terms zu kontrollieren.

Im Detail folgt

$$b(U, I_h^k(U^{p-1})) = b(U, U^{p-1}) - \left[b(U, U^{p-1}) - b(U, I_h^k(U^{p-1}))\right]$$

und analog zum Fall $p = 2$ gilt

$$b(U, U^{p-1}) = b_1(U, U^{p-1}) - \sum_{i=4}^{5} F_i(f, U^p/p, U, 1)$$

mit

$$b_1(U, U^{p-1}) = \frac{1}{p}\int_\Omega (U_-^N)^p\,dx_x - \frac{1}{p}\int_\Omega (u_0)^p\,dx_x + \sum_{i=0}^{5} E_i(f, U^p/p, U, 1),$$

so dass (5.63) lautet

$$\begin{aligned}
&\frac{1}{p}\int_\Omega \left(U_-^N\right)^p \, dx_x + \sum_{i=0}^{5} E_i(f, U^p/p, U, 1) - \left[b(U, U^{p-1}) - b(U, I_h^k(U^{p-1}))\right] \\
&+ \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \left(\hat{\epsilon}_2(U) + \hat{\epsilon}_3(U)\right) \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, I_h^k(U^{p-1})) \left(\nabla U, \nabla I_h^k(U^{p-1})\right)_{0,T} \qquad (5.64)\\
&= \frac{1}{p}\int_\Omega (u_0)^p \, dx_x + \sum_{i=4}^{5} F_i(f, U^p/p, U, 1).
\end{aligned}$$

Aufgrund der Ergebnisse des Abschnitts 5.3 für konvexe Entropiefunktionen η sind die Terme $E_1(f, U^p/p, U, 1)$, $E_2(f, U^p/p, U, 1) + E_3(f, U^p/p, U, 1)$ nichtnegativ. Analog zu (5.38) resultiert für E_4, E_5 :

$$\begin{aligned}
&E_4(f, U^p/p, U, 1) + E_5(f, U^p/p, U, 1) \\
&\geq \frac{1}{2} C_0^{\partial\Omega} \|U^+\|_{0,p,\Sigma_T}^p \geq 0.
\end{aligned}$$

Mit

$$\int_{R_{n+1}} \frac{1}{2}[f(U^-) - f(U^+)] \cdot n^+ + C_T(U^+ - U^-) \, dx_x = 0$$

folgt für den Interpolationsfehler

$$\begin{aligned}
&\left[b(U, U^{p-1}) - b(U, I_h^k(U^{p-1}))\right] \\
&\underset{(5.22)}{=} \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \Big\{ \int_T \nabla \cdot f(U) \left[U^{p-1} - I_h^k(U^{p-1})\right] \, dx \\
&+ \hat{\epsilon}_1(U) \Big(\hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, U^{p-1}) \left(\nabla U, \nabla U^{p-1}\right)_{0,T} - \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, I_h^k(U^{p-1})) \left(\nabla U, \nabla I_h^k(U^{p-1})\right)_{0,T} \Big) \\
&+ \int_{\partial^* T} \frac{1}{2} \left[f(U_T^-) - f(U_T^+)\right] \cdot n^+ \left[U^{p-1} - I_h^k(U^{p-1})\right] \, ds \\
&+ \int_{\partial^* T} C_T(U_T^+ - U_T^-) \left[U^{p-1} - I_h^k(U^{p-1})\right] \, ds \Big\} = \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h} \sum_{i=1}^{4} A_T^i.
\end{aligned} \qquad (5.65)$$

Die einzelnen Bestandteile des Interpolationsfehlers lassen sich mit den Lemmata 5.4.8 und 5.4.9 abschätzen:

$$\begin{aligned}
\left|A_T^1\right| &\leq \|(I - I_h^k)U^{p-1}\|_{0,\infty,T} \int_T |\nabla \cdot f(U)| \, dx \\
&\underset{(4.12)}{\leq} C h_T^{k+1} |U^{p-1}|_{k+1,\infty,T} \int_T |\nabla \cdot f(U)| \, dx \\
&\underset{(5.62)}{\leq} C p^{k+1} h_T^2 \|\nabla U\|_{0,\infty,T}^2 \|U\|_{0,\infty,T}^{p-3} \int_T |\nabla \cdot f(U)| \, dx
\end{aligned}$$

$$
\begin{aligned}
&= Cp^{k+1}h_T^2 \int_T \|\nabla U\|_{0,\infty,T}^2 \|U\|_{0,\infty,T}^{p-3} |\nabla \cdot f(U)|\, dx \\
&\leq Cp^{k+1}h_T^2 \int_{T\cap\{|U|>1\}} \|\nabla U\|_{0,\infty,T}^2 \|U\|_{0,\infty,T}^{p-2} |\nabla \cdot f(U)|\, dx \\
&\quad + Cp^{k+1}h_T^2 \int_{T\cap\{|U|\leq 1\}} \|\nabla U\|_{0,\infty,T}^2 |\nabla \cdot f(U)|\, dx \\
&\leq Cp^{k+1}h_T^2 \max_T(|\nabla \cdot f(U)|) \|U\|_{0,\infty,T}^{p-2} \int_T \|\nabla U\|_{0,\infty,T}^2\, dx \\
&\quad + Cp^{k+1}h_T^2 \max_T(|\nabla \cdot f(U)|) \int_T \|\nabla U\|_{0,\infty,T}^2\, dx.
\end{aligned}
$$

Wegen der lokalen Quasiuniformität der Familie $\{\mathcal{T}_h^n\}_{h>0}$ von affinen Partitionierungen gilt zusammen mit einer inversen Ungleichung für $v \in W_h$

$$
\begin{aligned}
\int_T \|\nabla v\|_{0,\infty,T}^2\, dx = |T| \|\nabla v\|_{0,\infty,T}^2 &\underset{(4.9)}{\leq} Ch_T^d \left(h_T^{-d/2} \|\nabla v\|_{0,2,T}\right)^2 \\
&= C \int_T \|\nabla v\|_{l^2}^2\, dx. \qquad (5.66)
\end{aligned}
$$

Insgesamt gilt mit

$$
(P_h'(\nabla U), P_h'(\nabla I_h^k(U^{p-1}))_{0,T,h} \underset{(5.59)}{\geq} 0 \underset{(5.58)}{\Rightarrow} \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, I_h^k(U^{p-1})) \geq \hat{\epsilon}_{\text{vms}}^{\min} > 0 \qquad (5.67)
$$

für den ersten Interpolationsfehler

$$
\begin{aligned}
|A_T^1| &\underset{\substack{(5.58),\\(5.66)}}{\leq} Cp^{k+1}h_T^2 \max_T(|\nabla \cdot f(U)|) \left\{ \int_T \nabla U \cdot \nabla I_h^k(U^{p-1})\, dx + \int_T \|\nabla U\|_{l^2}^2\, dx \right\} \\
&\leq Cp^{k+1} \frac{\hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, I_h^k(U^{p-1}))}{\hat{\epsilon}_{\text{vms}}^{\min}} h_T^2 \max_T(|\nabla \cdot f(U)|) \left(\nabla U, \nabla I_h^k(U^{p-1})\right)_{0,T} \\
&\quad + Cp^{k+1} \frac{\hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, U)}{\hat{\epsilon}_{\text{vms}}^{\min}} h_T^2 \max_T(|\nabla \cdot f(U)|) \|\nabla U\|_{0,2,T}^2
\end{aligned}
$$

und für den nächsten Term

$$
\begin{aligned}
\left|A_T^2\right| &\underset{(5.13)}{\leq} \mathcal{L}_{[s]}\hat{\epsilon}_1(U) \left| \left(P_h'(\nabla U), P_h'(\nabla [U^{p-1} - I_h^k(U^{p-1})])\right)_{0,T,h} \right| \\
&\underset{(4.52)}{\leq} \mathcal{L}_{[s]}\hat{\epsilon}_1(U) \sum_{i=n_{\text{dof}}^K+1}^{n_{\text{dof}}^k} \omega_i^J \left| (\nabla U)(x_i)(\nabla [U^{p-1} - I_h^k(U^{p-1})])(x_i) \right| \\
&\leq \mathcal{L}_{[s]}\hat{\epsilon}_1(U) \|\nabla U \cdot \nabla [U^{p-1} - I_h^k(U^{p-1})]\|_{0,1,T,h} \\
&\leq \mathcal{L}_{[s]}\hat{\epsilon}_1(U) \max_{x_i \in \mathcal{Q}} \left(\|(\nabla [U^{p-1} - I_h^k(U^{p-1})])(x_i)\|_{l^2} \right) \sum_{i=1}^{n_{\text{dof}}^k} \omega_i^J \|(\nabla U)(x_i)\|_{l^2} \\
&\underset{\substack{(2.6),\\(4.40)}}{\leq} C\sqrt{d}\hat{\epsilon}_1(U) \max_{x_i \in \mathcal{Q}} \left(\|(\nabla [U^{p-1} - I_h^k(U^{p-1})])(x_i)\|_{l^\infty} \right) \int_T \|\nabla U\|_{l^2}\, dx
\end{aligned}
$$

$$
\begin{aligned}
&\leq C\|\nabla[U^{p-1} - I_h^k(U^{p-1})]\|_{0,\infty,T} \int_T \hat{\epsilon}_1(U)\|\nabla U\|_{l^2}\,dx \\
&\underset{(4.12)}{\leq} Ch_T^k|U^{p-1}|_{k+1,\infty,T} \int_T \hat{\epsilon}_1(U)\|\nabla U\|_{l^2}\,dx \\
&\underset{(5.62)}{\leq} Cp^{k+1}h_T\|\nabla U\|_{0,\infty,T}^2\|U\|_{0,\infty,T}^{p-3} \int_T \hat{\epsilon}_1(U)\|\nabla U\|_{l^2}\,dx \\
&\leq Cp^{k+1}h_T^2\max_{x\in T}(\|f'(U)\|_{l^2})\max_{x\in T}(\|\nabla U\|_{l^2})\|U\|_{0,\infty,T}^{p-2} \int_T \|\nabla U\|_{l^2}^2\,dx \\
&+ Cp^{k+1}h_T^2\max_{x\in T}(\|f'(U)\|_{l^2})\max_{x\in T}(\|\nabla U\|_{l^2}) \int_T \|\nabla U\|_{l^2}^2\,dx \\
&\underset{(5.58)}{\leq} Cp^{k+1}\frac{\hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, I_h^k(U^{p-1}))}{\hat{\epsilon}_{\text{vms}}^{\min}}h_T^2\max_{x\in T}(\|f'(U)\|_{l^2})\max_{x\in T}(\|\nabla U\|_{l^2}) \left(\nabla U, \nabla I_h^k(U^{p-1})\right)_{0,T} \\
&+ Cp^{k+1}\frac{\hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, U)}{\hat{\epsilon}_{\text{vms}}^{\min}}h_T^2\max_{x\in T}(\|f'(U)\|_{l^2})\max_{x\in T}(\|\nabla U\|_{l^2})\|\nabla U\|_{0,2,T}^2.
\end{aligned}
$$

Die Randterme des Interpolationsfehlers werden wie folgt abgeschätzt:

$$
\begin{aligned}
|A_T^3| &\leq \|(I - I_h^k)U^{p-1}\|_{0,\infty,\partial^* T} \int_{\partial^* T} \frac{1}{2}\left|[f(U_T^-) - f(U_T^+)]\cdot n^+\right| ds \\
&\underset{(4.13)}{\leq} Ch_T^{k+1}|U^{p-1}|_{k+1,\infty,T} \int_{\partial^* T} \left|[f(U_T^-) - f(U_T^+)]\cdot n^+\right| ds \\
&\underset{(5.62)}{\leq} Cp^{k+1}h_T^2\|\nabla U\|_{0,\infty,T}^2\|U\|_{0,\infty,T}^{p-3} \int_{\partial^* T} \left|[f(U_T^-) - f(U_T^+)]\cdot n^+\right| ds \\
&\leq Cp^{k+1}h_T^2\|\nabla U\|_{0,\infty,T}^2\|U\|_{0,\infty,T}^{p-3}\max_{\partial^* T}(|[f(U_T^-) - f(U_T^+)]\cdot n^+|)|\partial T| \\
&\leq Cp^{k+1}h_T\|\nabla U\|_{0,2,T}^2\|U\|_{0,\infty,T}^{p-3}\max_{\partial^* T}(|[f(U_T^+) - f(U_T^-)]\cdot n^+|) \\
&\leq Cp^{k+1}h_T\max_{\partial^* T}(|[f(U_T^+) - f(U_T^-)]\cdot n^+|) \int_T \|\nabla U\|_{l^2}^2\|U\|_{0,\infty,T}^{p-3}\,dx \\
&\leq Cp^{k+1}h_T\max_{\partial^* T}(|[f(U_T^+) - f(U_T^-)]\cdot n^+|) \int_T \|\nabla U\|_{l^2}^2\|U\|_{0,\infty,T}^{p-2} + \|\nabla U\|_{l^2}^2\,dx \\
&\underset{(5.58)}{\leq} Cp^{k+1}h_T\max_{\partial^* T}(|[f(U_T^+) - f(U_T^-)]\cdot n^+|) \int_T \nabla U\cdot\nabla I_h^k(U^{p-1}) + \|\nabla U\|_{l^2}^2\,dx.
\end{aligned}
$$

Entsprechend ergibt sich

$$
|A_T^4| \leq Cp^{k+1}h_T\max_{\partial^* T}(C_T|U_T^+ - U_T^-|) \int_T \nabla U\cdot\nabla I_h^k(U^{p-1}) + \|\nabla U\|_{l^2}^2\,dx,
$$

so dass für den Gesamtinterpolationsfehler mit den an dieser Stelle gültigen Konstanten C_7

resultiert

$$
\begin{aligned}
&\left|\left[b(U,U^{p-1}) - b(U,I_h^k(U^{p-1}))\right]\right| \\
&\leq C_7 p^{k+1} \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h} h_T^2 \left\{ \max_{x\in T}(\|f'(U)\|_{l^2}) \max_{x\in T}(\|\nabla U\|_{l^2}) + R(U) \right\} \times \\
&\quad \times \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, I_h^k(U^{p-1})) \left(\nabla U, \nabla I_h^k(U^{p-1})\right)_{0,T} \\
&+ C_7 p^{k+1} \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h} h_T^2 \left\{ \max_{x\in T}(\|f'(U)\|_{l^2}) \max_{x\in T}(\|\nabla U\|_{l^2}) + R(U) \right\} \times \\
&\quad \times \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, U) \|\nabla U\|_{0,2,T}^2 \\
&\underset{\substack{(5.44),\\ C_3=C_2}}{\leq} C_7 p^{k+1} \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h} h_T^2 \left\{ \max_{x\in T}(\|f'(U)\|_{l^2}) \max_{x\in T}(\|\nabla U\|_{l^2}) + R(U) \right\} \times \\
&\quad \times \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, I_h^k(U^{p-1})) \left(\nabla U, \nabla I_h^k(U^{p-1})\right)_{0,T} \\
&+ \frac{C_7}{C_2} h_T^\beta p^{k+1} C_5.
\end{aligned}
\tag{5.68}
$$

(5.64) lässt sich somit schreiben als

$$
\begin{aligned}
&\int_\Omega \left(U_-^N\right)^p \, dx_x - C_7 p^{k+2} \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h} h_T^2 \left\{ \max_{x\in T}(\|f'(U)\|_{l^2}) \max_{x\in T}(\|\nabla U\|_{l^2}) + R(U) \right\} \times \\
&\qquad \times \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, I_h^k(U^{p-1})) \left(\nabla U, \nabla I_h^k(U^{p-1})\right)_{0,T} \\
&\qquad + p \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} (\hat{\epsilon}_2(U) + \hat{\epsilon}_3(U)) \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, I_h^k(U^{p-1})) \left(\nabla U, \nabla I_h^k(U^{p-1})\right)_{0,T} \\
&\qquad + p \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \hat{\epsilon}_1(U) \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, U^{p-1}) \left(\nabla U, \nabla (U^{p-1})\right)_{0,T} \\
&\qquad + \frac{p}{2} C_0^{\partial\Omega} \|U^+\|_{0,p,\Sigma_T}^p \\
&\qquad \leq \int_\Omega (u_0)^p \, dx_x + p \sum_{i=4}^{5} F_i(f, U^p/p, U, 1) + \frac{C_7}{C_2} h^\beta p^{k+2} C_5.
\end{aligned}
\tag{5.69}
$$

Wird $C_7 p^{k+2} \leq h^{-\beta}$, $0 \leq \beta \leq 1/2$ gewählt, so folgt

$$
\begin{aligned}
&\int_\Omega \left(U_-^N\right)^p dx_x + (p-1) \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} (\hat{\epsilon}_2(U) + \hat{\epsilon}_3(U)) \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, I_h^k(U^{p-1})) \left(\nabla U, \nabla I_h^k(U^{p-1})\right)_{0,T} \\
&\qquad + p \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \hat{\epsilon}_1(U) \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U, U^{p-1}) \left(\nabla U, \nabla (U^{p-1})\right)_{0,T} + \frac{p}{2} C_0^{\partial\Omega} \|U^+\|_{0,p,\Sigma_T}^p \\
&\qquad \leq \int_\Omega (u_0)^p \, dx_x + p \sum_{i=4}^{5} F_i(f, U^p/p, U, 1) + \frac{C_5}{C_2}.
\end{aligned}
\tag{5.70}
$$

Die Schranke $\beta \leq 1/2$ ist notwendig für die Konvergenz des Verfahrens (vgl. (6.19)). Da h hinreichend klein gewählt werden kann, bedeutet dies keine Einschränkung an p. Der Term

$$(p-1)\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}(\hat{\epsilon}_2+\hat{\epsilon}_3)\hat{\epsilon}_{\text{vms}}^{\text{coer}}(U,I_h^k(U^{p-1}))\left(\nabla U,\nabla I_h^k(U^{p-1})\right)_{0,T}$$

ist nach Defintion von p, $\hat{\epsilon}(U)$ und Lemma $5.4.8$ echt positiv. Die Terme F_4 und F_5 lassen sich analog zu $p=2$ behandeln und ergeben mit der Youngschen Ungleichung für den Parameter

$$\epsilon = \left(\frac{1}{2}\frac{p}{p-1}C_0^{\partial\Omega}\right)^{-(p-1)}$$

die Ungleichung

$$\frac{5}{4}pC_0^{\partial\Omega}\int_{\Sigma_T}|g_D||U^+|^{p-1}\,ds \leq \frac{1}{2}C_0^{\partial\Omega}\left(\frac{5}{2}\right)^p\left(\frac{p-1}{p}\right)^{p-1}\|g_D\|_{0,p,\Sigma_T}^p+\frac{p}{2}C_0^{\partial\Omega}\|U^+\|_{0,p,\Sigma_T}^p.$$

Also gilt

$$\begin{aligned}\int_\Omega\left(U_-^N\right)^p dx_x &+ (p-1)\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}(\hat{\epsilon}_2(U)+\hat{\epsilon}_3(U))\hat{\epsilon}_{\text{vms}}^{\text{coer}}(U,I_h^k(U^{p-1}))\left(\nabla U,\nabla I_h^k(U^{p-1})\right)_{0,T}\\ &+p\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\hat{\epsilon}_1(U)\hat{\epsilon}_{\text{vms}}^{\text{coer}}(U,U^{p-1})\left(\nabla U,\nabla(U^{p-1})\right)_{0,T}\\ &\leq\int_\Omega(u_0)^p\,dx_x+\frac{1}{2}C_0^{\partial\Omega}\left(\frac{5}{2}\right)^p\|g_D\|_{0,p,\Sigma_T}^p+\frac{C_5}{C_2}.\end{aligned}\tag{5.71}$$

Im nächsten Schritt ergibt sich nach (5.46) für $t_{N-1}\leq t\leq t_N$

$$\|U(t,\cdot)\|_{0,p,\Omega}^p\leq\|U_-^N\|_{0,p,\Omega}^p+p\int_t^{t_N}\int_\Omega|U(t',\cdot)|^{p-1}|\operatorname{div}f(U(t',\cdot))|\,dx_x dt'.$$

Aus der Hölderschen und Youngschen Ungleichung resultiert für

$$\begin{aligned}&p\int_t^{t_N}\int_\Omega|U(t',\cdot)|^{p-1}|\operatorname{div}f(U(t',\cdot))|\,dx_x dt'\\ &\quad\leq p\left(\int_t^{t_N}\int_\Omega|\operatorname{div}f(U)|^2|U|^{p-2}\,dx_x dt'\right)^{1/2}\left(\int_t^{t_N}\int_\Omega|U|^p\,dx_x dt'\right)^{1/2}\\ &\quad\leq\frac{p}{2}\left\{\int_t^{t_N}\int_\Omega\epsilon\,|\operatorname{div}f(U)|^2|U|^{p-2}\,dx_x dt'+\int_t^{t_N}\int_\Omega\frac{1}{\epsilon}|U|^p\,dx_x dt'\right\}\end{aligned}$$

mit

$$\begin{aligned}\frac{p}{2}\sum_{T\in\mathcal{T}_h^{N-1}}\epsilon\int_T|f'(U)\cdot\nabla U|^2|U|^{p-2}\,dx &\le \frac{p}{2}\sum_{T\in\mathcal{T}_h^{N-1}}\epsilon\max_{x\in T}\left(\|f'(U)\|_{l^2}^2\right)\int_T\|\nabla U\|_{l^2}^2|U|^{p-2}\,dx\\ &= p\sum_{T\in\mathcal{T}_h^{N-1}}\int_T\hat{\epsilon}_1(U)\|\nabla U\|_{l^2}^2(p-1)U^{p-2}\,dx\\ &= p\sum_{T\in\mathcal{T}_h^{N-1}}\int_T\hat{\epsilon}_1(U)\nabla U\cdot\nabla\left(U^{p-1}\right)\,dx\end{aligned}$$

und dem Parameter

$$\epsilon=\frac{2(p-1)C_1h_T}{\max\limits_{x\in T}\|f'(U)\|_{l^2}}$$

die Abschätzung

$$\begin{aligned}\|U(t,\cdot)\|_{0,p,\Omega}^p \le{}& \|U_-^N\|_{0,p,\Omega}^p\\ &+p\sum_{T\in\mathcal{T}_h^{N-1}}\int_T\hat{\epsilon}_1(U)\nabla U\cdot\nabla\left(U^{p-1}\right)\,dx\\ &+\frac{pC_0^\Omega}{4(p-1)C_1\min\limits_{T\in\mathcal{T}_h^{N-1}}\{h_T\}}\int_t^{t_N}\|U(t',\cdot)\|_{0,p,\Omega}^p\,dt'.\end{aligned}$$

Insgesamt liefert die letzte Ungleichung für $t_{N-1}\le t\le t_N$ mit der Gronwallschen Ungleichung und mit (5.71)

$$\begin{aligned}\|U(t,\cdot)\|_{0,p,\Omega}^p &\le C_6\left[p\sum_{T\in\mathcal{T}_h^{N-1}}\int_T\hat{\epsilon}_1(U)\nabla U\cdot\nabla\left(U^{p-1}\right)\,dx+\|U_-^N\|_{0,p,\Omega}^p\right]\\ &\le \frac{C_6}{\hat{\epsilon}_{\text{vms}}^{\min}}\left[p\sum_{T\in\mathcal{T}_h^{N-1}}\int_T\hat{\epsilon}_1(U)\hat{\epsilon}_{\text{vms}}^{\text{coer}}(U,U^{p-1})\nabla U\cdot\nabla\left(U^{p-1}\right)\,dx+\|U_-^N\|_{0,p,\Omega}^p\right]\\ &\le \frac{C_6}{\hat{\epsilon}_{\text{vms}}^{\min}}\left[\|u_0\|_{0,p,\Omega}^p+\frac{1}{2}C_0^{\partial\Omega}\left(\frac{5}{2}\right)^p\|g_D\|_{0,p,\Sigma_T}^p+\frac{C_5}{C_2}\right],\end{aligned}\tag{5.72}$$

wobei

$$\exp(\frac{pC_0^\Omega}{4(p-1)C_1\min\limits_{T\in\mathcal{T}_h^{N-1}}\{h_T\}}(t_N-t))<C_6$$

gesetzt wird (vgl. (5.48)). Dies gewährleistet die Gültigkeit der Ungleichung

$$\sup_{t\ge0}\|U(t,\cdot)\|_{0,p,\Omega}\le\left(\frac{C_6}{\hat{\epsilon}_{\text{vms}}^{\min}}\right)^{1/p}\left[\|u_0\|_{0,p,\Omega}+\frac{5}{2}\left(\frac{1}{2}C_0^{\partial\Omega}\right)^{1/p}\|g_D\|_{0,p,\Sigma_T}+\left(\frac{C_5}{C_2}\right)^{1/p}\right]\tag{5.73}$$

für $4 \leq p \leq C_7^{-\frac{1}{k+2}} h^{-\frac{\beta}{k+2}}$. Die Beschränktheit von $\|U\|_{0,\infty,Q_T}$ ergibt sich aus der Anwendung der inversen Ungleichung (4.69)

$$\begin{aligned}
\|U\|_{0,\infty,Q_T} &\leq \left(\frac{2(p+1)k^2}{C_{qu}h}\right)^{\frac{d_x}{p}} \sup_{t\geq 0}\|U(t,\cdot)\|_{0,p,\Omega} \\
&\leq \left(\frac{5k^2}{2C_{qu}h^{1+\frac{\beta}{k+2}}}\right)^{\frac{d_x}{p}} \sup_{t\geq 0}\|U(t,\cdot)\|_{0,p,\Omega} \\
&= \left(\frac{5k^2}{2C_{qu}}\right)^{\frac{d_x}{p}} \exp\left(\left(1+\frac{\beta}{k+2}\right)\frac{d_x}{p}\ln\left(\frac{1}{h}\right)\right) \sup_{t\geq 0}\|U(t,\cdot)\|_{0,p,\Omega}.
\end{aligned}$$

Ferner gilt speziell für $\beta_0 = \beta/(k+2)$, $p = C_7^{-\frac{1}{k+2}} h^{-\beta_0}$ und $h \leq 1$

$$\|U\|_{0,\infty,Q_T} \leq C_8 \sup_{t\geq 0}\|U(t,\cdot)\|_{0,p,\Omega} \tag{5.74}$$

mit

$$C_8 = C_8(h) = \left(\frac{5k^2}{2C_{qu}}\right)^{C_7^{\frac{1}{k+2}} d_x h^{\beta_0}} \exp\left((1+\beta_0)\, C_7^{\frac{1}{k+2}} d_x h^{\beta_0} \ln\left(\frac{1}{h}\right)\right).$$

Damit ist folgendes Theorem gezeigt:

Theorem 5.4.10 *Es sei Ω ein Gebiet mit Lipschitz-Rand und $\{\mathcal{T}_h\}_{h>0}$ eine quasiuniforme Familie von Partitionierungen von $(0,T)\times\Omega$. Dann existiert mit $C_2 = C_3 = C_5$ ein C unabhängig von h, so dass für die Lösung $U \in W_h$ von (5.15) zusammen mit der Annahme (5.8) und (5.9) die Ungleichung*

$$\begin{aligned}
\|U\|_{0,\infty,Q_T} \leq \max_{0<h\leq 1} C_8(h) \max\{\frac{C_6}{\hat{\epsilon}^{\min}_{\text{vms}}}, 1\} \Bigg[&\max\{|\Omega|,1\}\|u_0\|_{0,\infty,\Omega} \\
&+ \max\{|\Sigma_T|,1\}\frac{5}{2}\max\{\frac{1}{2}C_0^{\partial\Omega},1\}\|g_D\|_{0,\infty,\Sigma_T} + 1\Bigg] \leq C.
\end{aligned}$$

gilt.

Bemerkung 5.4.11 Die Aussage des vorstehenden Theorems kann mit den hier präsentierten Methoden im VMS-Kontext nur erreicht werden, sofern der verwendete Fluktuationsoperator die Eigenschaften

$$\begin{aligned}
\left(P_h'(\nabla v), P_h'(\nabla v^{p-1})\right)_{0,T,h} &\geq 0, \\
\left(P_h'(\nabla v), P_h'(\nabla I_h^k(v^{p-1}))\right)_{0,T,h} &\geq 0,
\end{aligned}$$

für $v \in \mathbb{Q}_k(T)$ und $p = 2m$, $m \in \mathbb{N}$ besitzt. Die erste Ungleichung motiviert die Konstruktion von P_h', während die zweite unkritisch ist, da $(P_h'(\nabla\varphi_i), P_h'(\nabla\varphi_j))_{0,T,h}$, $1 \leq i,j \leq n_{\text{dof}}^k$ symmetrisch, positiv semidefinit ist (siehe Lemma 5.4.8).

Theorem 5.4.12 *Unter den Voraussetzungen des Theorems 5.4.10 gilt mit $h \to 0$*

$$\|U\|_{0,\infty,Q_T} \leq \|u_0\|_{0,\infty,\Omega} + \frac{5}{2}\|g_D\|_{0,\infty,\Sigma_T} + 1. \tag{5.75}$$

Beweis Folgt mit $h = C_7^{-\frac{1}{\beta}} p^{-1/\beta_0} \to 0$ direkt aus (5.73) und (5.74). □

Bemerkung 5.4.13 In [Sze91] wird ebenfalls der Fall $k > 1$ für die Abschätzung von $\|U\|_{0,\infty,Q_T}$ behandelt. Da allerdings die Abschätzung (5.58) bisher nur für den Fall $k = 1$ bewiesen wurde, wird dort der Shock-capturing Term realisiert, indem er auf eine entsprechend feinere Triangulierung mit linearen Polynomen interpoliert wird. Diese Tatsache hat zur Folge, dass mit angepassten β eine Schranke der Form $C^p p^{k+2} \leq h^{-\beta}$ entsteht (vgl. [Sze91, (3.16)]), die erfüllt ist, falls $p \leq C \ln(1/h)$ gilt. Somit gilt

$$\lim_{h\to 0} C_8(h) \neq 1$$

und ein zu (5.75) vergleichbares Ergebnis kann nicht erzielt werden.

Theorem 5.4.14 *Unter den Voraussetzungen des Theorems 5.4.10 gilt mit* $k = 0$

$$\|U\|_{0,\infty,Q_T} \leq \|u_0\|_{0,\infty,\Omega} + \frac{5}{2}\|g_D\|_{0,\infty,\Sigma_T}, \quad \forall h > 0. \tag{5.76}$$

Beweis Mit $k = 0$ folgt $\left[b(U, U^{p-1}) - b(U, I_h^k(U^{p-1}))\right] = 0$ und $|U^{p-1}|_{k+1,\infty,T} = 0$. Aus diesem Grund besteht keine Notwendigkeit für die Beschränkung $C_7 p^{k+2} \leq h^{-\beta}$, so dass für $p \to \infty$ mit (5.73) und (5.74) die Behauptung folgt. □

6 Konvergenzanalysis

Im ersten Teil dieses Kapitels wird untersucht, ob die in der Einleitung angedeutete höhere Konvergenzordnung durch Einführung einer lokalen Projektion mit Hilfe des Fluktuationsoperators im nichtlinearen Shock-capturing Term erreicht werden kann. Zum Vergleich wird anschließend die residualbasierte Methode aus [JJS95], die einen zusätzlichen Stromliniendiffusionsterm besitzt, analysiert. Es ist anzumerken, dass keine a priori Abschätzung der bekannten Methode, weder bzgl. h noch bzgl. hk, in der Literatur zu finden ist. Allerdings wird in [Sze89a, Chapter 7] eine a priori Untersuchung bzgl. h einer Standard-Galerkin-Methode mit Stromliniendiffusionsterm und Shock-capturing vorgenommen. Diese ist jedoch auf die residuale Struktur der Methode angewiesen.

Um bei den Fehlerabschätzungen eine explizite Abhängigkeit des Polynomgrades zu erreichen, werden Lagrangesche Finite-Elemente bzgl. der Gauss-Lobatto Quadraturpunkte verwendet.

6.1 A priori Fehlerabschätzung I

In den vorherigen Kapiteln wurde die schwachen Formulierung (5.15) mit dem Entropie-Flux Paar η, q ausgestattet (vgl. (5.22) und (5.24)). Ohne die Verwendung des Entropie-Flux Paares, oder aber mit $\eta(v) = v^2/2$ und $\varphi = 1$, folgt für die Aufgabe (5.15) mit (5.22) und (5.26) sofort

Finde $U \in W_h$ derart, dass gilt

$$\begin{aligned} b_1(U,v) + \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_2(U) \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U,v) \left(\nabla U, \nabla v\right)_{0,T} \\ + \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_3(U) \hat{\epsilon}_{\text{vms}}^{\text{coer}}(U,v) \left(\nabla U, \nabla v\right)_{0,T} = l(v), \quad \forall v \in W_h. \end{aligned} \tag{6.1}$$

In diesem Fall ist

$$\begin{aligned} b_1(v,w) = &\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\int_T \nabla\cdot f(v)w\,dx + \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\hat{\epsilon}_1(v)\hat{\epsilon}_{\text{vms}}^{\text{coer}}(v,w)\,(\nabla v,\nabla w)_{0,T}\\ &+\sum_{n=1}^{N-1}\int_\Omega (v_+^n - v_-^n)w_+^n\,dx_x + \int_\Omega v_+^0 w_+^0\,dx_x\\ &+\sum_{n=0}^{N-1}\int_{R_{n,n+1}^i}\frac{1}{2}\left(f(v^-)-f(v^+)\right)\cdot n^+(w^+ + w^-)\,ds\\ &+\sum_{n=0}^{N-1}\int_{R_{n,n+1}^i} C_0^\Omega(v^+ - v^-)(w^+ - w^-)\,ds\\ &-\sum_{n=0}^{N-1}\int_{\Lambda_{n,n+1}^-}\frac{1}{2}f(v^+)\cdot n^+ w^+\,ds + \sum_{n=0}^{N-1}\int_{\Lambda_{n,n+1}^-} C_0^{\partial\Omega}v^+w^+\,ds \end{aligned} \tag{6.2}$$

und

$$l(v) = \int_\Omega u_0 v_+^0\,dx_x - \sum_{n=0}^{N-1}\int_{\Lambda_{n,n+1}^-}\frac{1}{2}f(g_D)\cdot n^+ v^+ + C_0^{\partial\Omega}g_D v^+\,ds. \tag{6.3}$$

Wie bereits angedeutet, geht mit (5.14) auf Elementen für die

$$-C_4/2 \le \hat{\epsilon}_{\text{vms}}(U,v) \le 5/2C_4$$

mit $v \in W_h$ gilt, der VMS-Charakter verloren, und es resultiert im günstigsten Fall ein Verfahren mit einer Fehlerordnung $\mathcal{O}(h)$. Für die nachfolgende Analysis sei also $s(x) = x$ bzw.

$$\hat{\epsilon}_{\text{vms}}(v,w) = \hat{\epsilon}_{\text{vms}}^{\text{coer}}(v,w), \quad \forall v,w \in W. \tag{6.4}$$

Die Konvergenzuntersuchung erfolgt für das lineare Problem mit

$$f(v) = bv, \quad b \in C(\overline{Q}_T)^d, \quad b_0 = 1, \quad \nabla\cdot b \in L^\infty(Q_T), \tag{6.5}$$

so dass es mit der Annahme

$$-\frac{1}{2}\nabla\cdot b(x) \ge \mu_0 > 0, \text{fast überall in } Q_T \tag{6.6}$$

möglich ist, eine dem Problem angepasste Norm zu definieren. Dies geschieht mit Hilfe der Identität

$$\begin{aligned} \int_{Q_T}\nabla\cdot f(v)v\,dx &= \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\frac{1}{2}\int_{\partial T} b\cdot n^+(v^+)^2\,ds - \frac{1}{2}\int_T(\nabla\cdot b)\,v^2\,dx\\ &= \frac{1}{2}\int_\Omega (v_-^N)^2\,dx_x + \sum_{n=1}^{N-1}\frac{1}{2}\int_\Omega (v_-^n)^2 - (v_+^n)^2\,dx_x - \frac{1}{2}\int_\Omega (v_+^0)^2\,dx_x\\ &\quad+\frac{1}{2}\sum_{n=0}^{N-1}\int_{R_{n,n+1}^i} b_x\cdot {n_x}^+((v^+)^2 - (v^-)^2)\,ds\\ &\quad+\frac{1}{2}\sum_{n=0}^{N-1}\int_{\Lambda_{n,n+1}} b_x\cdot {n_x}^+(v^+)^2\,ds - \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\frac{1}{2}\int_T(\nabla\cdot b)\,v^2\,dx, \end{aligned}$$

so dass eine gitterabhängige Norm via

$$\begin{aligned} b_1(v,v) &= C_1 \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \frac{h_T}{k} \max_{x\in T}(\|b\|_{l^2}) \|P_h'(\nabla v)\|_{0,2,T,h}^2 + \frac{1}{2}\|(-\nabla\cdot b)^{1/2} v\|_{0,2,Q_T}^2 \\ &+ \frac{1}{2}\left\{ \|v_-^N\|_{0,2,\Omega}^2 + \sum_{n=1}^{N-1} \|v_+^n - v_-^n\|_{0,2,\Omega}^2 + \|v_+^0\|_{0,2,\Omega}^2 \right\} \\ &+ C_0^{\Omega} \sum_{n=0}^{N-1} \|v^+ - v^-\|_{0,2,R_{n,n+1}^i}^2 + C_0^{\partial\Omega} \sum_{n=0}^{N-1} \|v^+\|_{0,2,\Lambda_{n,n+1}^-}^2 \\ &+ \frac{1}{2} \sum_{n=0}^{N-1} \|(b_x \cdot {n_x}^+)^{1/2} v^+\|_{0,2,\Lambda_{n,n+1}^+}^2 \\ &= |||v|||_h^2, \quad \forall v \in W_h \end{aligned} \tag{6.7}$$

definiert werden kann.

Ausgangspunkt der Untersuchung sind wie üblich zwei schwache Formulierungen, die jeweils die kontinuierliche und diskrete Lösung enthalten. Die Schwierigkeiten ergeben sich zum einen aus der nichtresidualen Struktur der Methode und zum anderen aus dem nichtlinearen Shockcapturing Term. Eine *Galerkin-Orthogonalität* kann daher nicht erwartet werden.

Sei $u \in W = W^{1,\infty}(Q_T) \subset C(\overline{Q}_T)$ eine schwache Lösung nach Definition 3.2.2. Die anschließende Bemerkung liefert zusammen mit (5.10) den Nachweis, dass u auch

$$a_1(u,v) = l(v) \quad \forall v \in W^{1,2}(Q_T, \mathcal{T}_h) \tag{6.8}$$

und

$$a_1(u,v) = l(v) \quad \forall v \in W_h, \tag{6.9}$$

bzw.

$$b_1(u,v) - \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \hat{\epsilon}_1 \left(P_h'(\nabla u), P_h'(\nabla v)\right)_{0,T,h} = l(v) \quad \forall v \in W_h, \tag{6.10}$$

erfüllt. Aus der Differenz der Gleichungen (6.1) und (6.10) folgt die Aussage

$$\begin{aligned} b_1(u-U,v) &- \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \hat{\epsilon}_1 \left(P_h'(\nabla u), P_h'(\nabla v)\right)_{0,T,h} \\ &- \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \hat{\epsilon}_2(U) \left(P_h'(\nabla U), P_h'(\nabla v)\right)_{0,T,h} \\ &- \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \hat{\epsilon}_3(U) \left(P_h'(\nabla U), P_h'(\nabla v)\right)_{0,T,h} = 0, \quad \forall v \in W_h. \end{aligned} \tag{6.11}$$

Mit den Notationen $e_h^k = U - I_{GL}^k(u)$, $e = u - U$ und $\xi = u - I_{GL}^k(u)$ folgt

$$b_1(e_h^k, e_h^k) + \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \hat{\epsilon}_2(U) \|P_h'(\nabla e_h^k)\|_{0,2,T,h}^2 + \hat{\epsilon}_3(U) \|P_h'(\nabla e_h^k)\|_{0,2,T,h}^2$$

$$
\begin{aligned}
&= b_1(\xi, e_h^k) - b_1(e, e_h^k) + \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_2(U) \left(P_h'(\nabla U) - P_h'(\nabla I_{GL}^k(u)), P_h'(\nabla e_h^k) \right)_{0,T,h} \\
&\quad + \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_3(U) \left(P_h'(\nabla U) - P_h'(\nabla I_{GL}^k(u)), P_h'(\nabla e_h^k) \right)_{0,T,h} \\
&\underset{(6.11)}{=} b_1(\xi, e_h^k) - \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_1 \left(P_h'(\nabla u), P_h'(\nabla e_h^k) \right)_{0,T,h} \\
&\quad - \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_2(U) \left(P_h'(\nabla I_{GL}^k(u)), P_h'(\nabla e_h^k) \right)_{0,T,h} \\
&\quad - \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_3(U) \left(P_h'(\nabla I_{GL}^k(u)), P_h'(\nabla e_h^k) \right)_{0,T,h} = \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV}
\end{aligned}
$$

bzw.

$$b_1(e_h^k, e_h^k) \leq b_1(\xi, e_h^k) + |\mathrm{II}| + |\mathrm{III}| + |\mathrm{IV}|.$$

Für die abschließende Behandlung der Terme $\mathrm{I}-\mathrm{IV}$ wird aufgrund des diskreten Innenproduktes $(\cdot,\cdot)_{0,T,h}$ bzw. $(\cdot,\cdot)_{0,T,GL}$ häufig die Eigenschaft (4.44) des Lumpingoperators genutzt. Dabei wechselt die Bezeichnung von $(\cdot,\cdot)_{0,T,h}$ zu $(\cdot,\cdot)_{0,T,GL}$ sobald die spezielle Lage der Gauss-Lobatto Punkte von Bedeutung wird.

Bei der Abschätzung von I liefert die Stetigkeit der Lösung u die Identitäten $\xi_+^n = \xi_-^n$ und $\xi^+ = \xi^-$. Letzteres bedeutet für (6.2)

$$
\begin{aligned}
b_1(\xi, e_h^k) = &\sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \int_T \nabla \cdot (b\xi) e_h^k \, dx + \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_1(\xi) \left(P_h'(\nabla \xi), P_h'(\nabla e_h^k) \right)_{0,T,h} \\
&+ \int_\Omega \xi_+^0 {e_h}_+^0 \, dx_x - \sum_{n=0}^{N-1} \int_{\Lambda_{n,n+1}^-} \frac{1}{2} b_x \cdot {n_x}^+ \xi^+ e_h^+ \, ds + \sum_{n=0}^{N-1} \int_{\Lambda_{n,n+1}^-} C_0^{\partial\Omega} \xi^+ e_h^+ \, ds.
\end{aligned}
$$

Um die Notation auf den Kanten nicht zu überfrachten, werden die Indizes h und k bei e_h^k vernachlässigt. Genau wie zuvor folgt für den ersten Term nach partieller Integration die Darstellung

$$
\begin{aligned}
&\sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \int_T \nabla \cdot (b\xi) e_h^k \, dx \\
&= \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \frac{1}{2} \int_{\partial T} b \cdot n^+ \xi^+ e_h^+ \, ds - \frac{1}{2} \int_T (\nabla \cdot b) \, \xi e_h^k \, dx \\
&= \frac{1}{2} \left\{ \int_\Omega \xi^N e_-^N \, dx_x + \sum_{n=1}^{N-1} \int_\Omega \xi^n (e_-^n - e_+^n) \, dx_x - \int_\Omega \xi^0 e_+^0 \, dx_x \right\} \\
&\quad + \frac{1}{2} \sum_{n=0}^{N-1} \int_{R_{n,n+1}^i} b_x \cdot {n_x}^+ \xi (e^+ - e^-) \, ds
\end{aligned}
$$

$$+\frac{1}{2}\sum_{n=0}^{N-1}\int_{\Lambda_{n,n+1}} b_x\cdot {n_x}^+\xi e^+\,ds-\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\frac{1}{2}\int_T(\nabla\cdot b)\,\xi e_h^k\,dx$$

und somit

$$\begin{aligned}
|\mathrm{I}| \le{}& \frac{1}{4}\left\{\|\xi^N\|_{0,2,\Omega}^2+\|e_-^N\|_{0,2,\Omega}^2+\sum_{n=1}^{N-1}\|\xi^n\|_{0,2,\Omega}^2+\sum_{n=1}^{N-1}\|e_+^n-e_-^n\|_{0,2,\Omega}^2\right\}\\
&+\frac{1}{4}C_0^{\Omega}\sum_{n=0}^{N-1}\left\{\|\xi\|_{0,2,R_{n,n+1}^i}^2+\|e^+-e^-\|_{0,2,R_{n,n+1}^i}^2\right\}\\
&+\frac{1}{4}C_0^{\partial\Omega}\sum_{n=0}^{N-1}\left\{\|\xi\|_{0,2,\Lambda_{n,n+1}^+}^2+\underbrace{\|e^+\|_{0,2,\Lambda_{n,n+1}^+}^2}_{=0}\right\}\\
&+\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\hat{\epsilon}_1\left(\|P_h'(\nabla\xi)\|_{0,2,T,h}^2+\frac{1}{4}\|P_h'(\nabla e_h^k)\|_{0,2,T,h}^2\right)\\
&+\frac{1}{2}C_0^{\partial\Omega}\sum_{n=0}^{N-1}\left(\|\xi\|_{0,2,\Lambda_{n,n+1}^-}^2+\|e^+\|_{0,2,\Lambda_{n,n+1}^-}^2\right)\\
&+\frac{1}{4}\|(-\nabla\cdot b)^{1/2}\xi\|_{0,2,Q_T}^2+\frac{1}{4}\|(-\nabla\cdot b)^{1/2}e_h^k\|_{0,2,Q_T}^2\\
\le{}&\frac{1}{4}b_1(e_h^k,e_h^k)+\frac{1}{4}\sum_{n=1}^{N}\|\xi^n\|_{0,2,\Omega}^2+\frac{1}{4}C_0^{\Omega}\sum_{n=0}^{N-1}\|\xi\|_{0,2,R_{n,n+1}^i}^2\\
&+\frac{1}{2}C_0^{\partial\Omega}\sum_{n=0}^{N-1}\|\xi\|_{0,2,\Lambda_{n,n+1}}^2+\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\hat{\epsilon}_1\|P_h'(\nabla\xi)\|_{0,2,T,h}^2\\
&+\frac{1}{4}\|(-\nabla\cdot b)^{1/2}\xi\|_{0,2,Q_T}^2\\
\le{}&\frac{1}{4}b_1(e_h^k,e_h^k)+\max\left(\frac{1}{4},\frac{1}{2}C_0^{\Omega},\frac{1}{2}C_0^{\partial\Omega}\right)\sum_{n=1}^{N}\sum_{T\in\mathcal{T}_h^n}\|\xi\|_{0,2,\partial T}^2\\
&+\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\hat{\epsilon}_1\|P_h'(\nabla\xi)\|_{0,2,T,h}^2+\frac{1}{4}\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\max_{x\in T}(-\nabla\cdot b)\|\xi\|_{0,2,T}^2.
\end{aligned}$$

Ferner gilt aufgrund der Definition (4.50) des Fluktuationsoperators P_h' :

$$\begin{aligned}
\|P_h'(\nabla\xi)\|_{0,2,T,h} &\le \|\nabla\xi\|_{0,2,T,h}\\
&\underset{\substack{\text{„}h=GL\text{“},\\(4.44)}}{\le} C\|\nabla\xi\|_{0,2,T}\\
&\underset{(4.55)}{\le} C_9\left(\frac{h_T}{k}\right)^k|u|_{k+1,2,T},\ C_9>0 \qquad (6.12)
\end{aligned}$$

und insgesamt

$$|\mathrm{I}| \le \frac{1}{4}b_1(e_h^k,e_h^k)+C\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\left(k^{\epsilon}\frac{h_T^{2k+1}}{k^{2k+1}}+\hat{\epsilon}_1\frac{h_T^{2k}}{k^{2k}}+\frac{h_T^{2k+2}}{k^{2k+2}}\right)|u|_{k+1,2,T}^2$$

$$\underset{(5.18)}{\leq} \frac{1}{4} b_1(e_h^k, e_h^k) + C \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} k^\epsilon \frac{h_T^{2k+1}}{k^{2k+1}} |u|_{k+1,2,T}^2, \quad 0 < \epsilon \ll 1.$$

Für die Abschätzung der nächsten Terme erweist sich $P_h'(\nabla I_{GL}^k(u))$, $P_h'(e_h^k) \in \mathbb{Q}_k(T)$ sowie die Zwischenüberlegungen

$$\begin{aligned}
\|P_h'(\nabla I_{GL}^k(u))\|_{0,2,T} &\underset{(4.57)}{\leq} C \left(\frac{h_T}{K}\right)^{K+1} \left\{ |\nabla I_{GL}^k(u)|_{K+1,2,T} + |P_h^{K,k}(\nabla I_{GL}^k(u))|_{K+1,2,T} \right\} \\
&\leq C \left(\frac{h_T}{K}\right)^{K+1} \left\{ |I_{GL}^k(u)|_{K+2,2,T} + |P_h^{K,k}(\nabla I_{GL}^k(u))|_{K+1,2,T} \right\} \\
&\leq C \left(\frac{h_T}{K}\right)^{K+1} \left\{ |u|_{K+2,2,T} + |u - I_{GL}^k(u)|_{K+2,2,T} \right. \\
&\qquad \left. + |P_h^{K,k}(\nabla I_{GL}^k(u))|_{K+1,2,T} \right\} \\
&\underset{(4.55)}{\leq} C_{10} \left(\frac{h_T}{K}\right)^{K+1} \left\{ |u|_{K+2,2,T} + |P_h^{K,k}(\nabla I_{GL}^k(u))|_{K+1,2,T} \right\} \qquad (6.13)
\end{aligned}$$

und

$$\begin{aligned}
\|\nabla I_{GL}^k(u)\|_{0,\infty,T} &\leq \| \sum_{i=1}^{n_{dof}} u_i^{(1)} \varphi_i \|_{0,\infty,T} \\
&\leq \sum_{i=1}^{n_{dof}} |u_i^{(1)}| \|\varphi_i\|_{0,\infty,T} \underset{(4.45)}{\leq} C_{11} \|u_T^{(1)}\|_{l^1} \qquad (6.14)
\end{aligned}$$

mit $C_{10}, C_{11} > 0$ als hilfreich.

$$\begin{aligned}
|\mathrm{II}| &\leq \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_1 \left| \left(P_h'(\nabla u), P_h'(\nabla e_h^k)\right)_{0,T,h} \right| \\
&\leq \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_1 \|P_h'(\nabla u)\|_{0,2,T,h} \|P_h'(\nabla e_h^k)\|_{0,2,T,h} \\
&\leq \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_1 \left\{ \|P_h'(\nabla u)\|_{0,2,T,h}^2 + \frac{1}{4} \|P_h'(\nabla e_h^k)\|_{0,2,T,h}^2 \right\} \\
&\leq \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} 2\hat{\epsilon}_1 \left\{ \|P_h'(\nabla I_{GL}^k(u))\|_{0,2,T,h}^2 + \|P_h'(\nabla u - \nabla I_{GL}^k(u))\|_{0,2,T,h}^2 \right\} \\
&\qquad + \frac{1}{4} b_1(e_h^k, e_h^k) \\
&\underset{\substack{\text{„}h=GL\text{“}, \\ (6.12)}}{\leq} \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} 2\hat{\epsilon}_1 \left\{ \|P_{GL}'(\nabla I_{GL}^k(u))\|_{0,2,T,GL}^2 + C_9 \left(\frac{h_T}{k}\right)^{2k} |u|_{k+1,2,T}^2 \right\}
\end{aligned}$$

$$
\begin{aligned}
&\quad + \frac{1}{4} b_1(e_h^k, e_h^k) \\
&\underset{(6.13)}{\overset{(4.44)}{\leq}} \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} C \frac{h_T}{k} \left(\frac{h_T}{K} \right)^{2(K+1)} \left\{ |u|^2_{K+2,2,T} + |P_{GL}^{K,k}(\nabla I_{GL}^k(u))|^2_{K+1,2,T} \right\} \\
&\quad + C \left(\frac{h_T}{k} \right)^{2k+1} |u|^2_{k+1,2,T} + \frac{1}{4} b_1(e_h^k, e_h^k),
\end{aligned}
$$

$$
\begin{aligned}
|\mathrm{III}| \quad &\leq \quad \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_2 \left| \left(P_h'(\nabla I_{GL}^k(u)), P_h'(\nabla e_h^k) \right)_{0,T,h} \right| \\
&\underset{(4.44)}{\overset{\text{„}h=GL\text{“},}{\leq}} \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} C_{12} \hat{\epsilon}_2 \| P_{GL}'(\nabla I_{GL}^k(u)) \|_{0,2,T,GL} \| P_{GL}'(\nabla e_h^k) \|_{0,2,T} \\
&\leq \quad \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} C_{12} C_2 \left(\frac{h_T}{k^{2+d/2}} \right)^{2-\beta} \max_{x \in T}(\|b\|_{l^2}) \max_{x \in T}(\|\nabla U\|_{l^2}) \times \\
&\quad \times \| P_{GL}'(\nabla I_{GL}^k(u)) \|_{0,2,T} \| P_{GL}'(\nabla e_h^k) \|_{0,2,T} \\
&\leq \quad \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} C_{12} C_2^2 \frac{\max_{x \in T}(\|b\|_{l^2})^2 \max_{x \in T}(\|\nabla U\|_{l^2})^2 k}{C_{13} 2 \max_{x \in T}(\|b\|_{l^2}) h_T} \left(\frac{h_T}{k^{2+d/2}} \right)^{2(2-\beta)} \| P_{GL}'(\nabla I_{GL}^k(u)) \|^2_{0,2,T,GL} \\
&\quad + \quad C_{12} \frac{C_{13} \max_{x \in T}(\|b\|_{l^2}) h_T}{2k} \| P_{GL}'(\nabla e_h^k) \|^2_{0,2,T} \\
&\leq \quad \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} C_{12} C_2^2 \frac{\max_{x \in T}(\|b\|_{l^2}) \max_{x \in T}(\|\nabla e_h^k\|_{l^2})^2 k}{C_{13} h_T} \left(\frac{h_T}{k^{2+d/2}} \right)^{2(2-\beta)} \| P_{GL}'(\nabla I_{GL}^k(u)) \|^2_{0,2,T,GL} \\
&\quad + \quad C_{12} C_2^2 \frac{\max_{x \in T}(\|b\|_{l^2}) \max_{x \in T}(\|\nabla I_{GL}^k(u)\|_{l^2})^2 k}{C_{13} h_T} \left(\frac{h_T}{k^{2+d/2}} \right)^{2(2-\beta)} \| P_{GL}'(\nabla I_{GL}^k(u)) \|^2_{0,2,T,GL} \\
&\quad + \quad C_{12} \frac{C_{13} \max_{x \in T}(\|b\|_{l^2}) h_T}{2k} \| P_{GL}'(\nabla e_h^k) \|^2_{0,2,T} \\
&\underset{(6.13)}{\overset{(6.14)}{\leq}} \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} C_{12} C_2^2 \frac{\max_{x \in T}(\|b\|_{l^2}) d |e_h^k|^2_{1,\infty,T}}{C_{13}} \frac{k}{h_T} \left(\frac{h_T}{k^{2+d/2}} \right)^{2(2-\beta)} \| \nabla I_{GL}^k(u) \|^2_{0,2,T,GL} \\
&\quad + \quad C_{10}^2 C_{11}^2 C_{12} C_2^2 \frac{d \max_{x \in T}(\|b\|_{l^2})}{C_{13}} \frac{k}{h_T} \left(\frac{h_T}{k^{2+d/2}} \right)^{2(2-\beta)} \left(\frac{h_T}{K} \right)^{2(K+1)} \left\{ \|u_T^{(1)}\|^2_{l^1} |u|^2_{K+2,2,T} \right. \\
&\quad + \quad \left. |P_{GL}^{K,k}(\nabla I_{GL}^k(u))|^2_{K+1,2,T} \right\} + C_{12} \frac{C_{13} \max_{x \in T}(\|b\|_{l^2}) h_T}{2k} \| P_{GL}'(\nabla e_h^k) \|^2_{0,2,T} \\
&\underset{(6.14)}{\overset{(4.61)}{\leq}} \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} C_{11}^2 C_{12} C_{14}^2 C_2^2 \frac{d \max_{x \in T}(\|b\|_{l^2})}{C_{13}} \frac{h_T^{1-2\beta}}{k^{3-\beta(4+d)}} \|u_T^{(1)}\|^2_{l^1} \|e_h^k\|^2_{0,2,T} \\
&\quad + \quad C_{10}^2 C_{11}^2 C_{12} C_2^2 \frac{d \max_{x \in T}(\|b\|_{l^2})}{C_{13}} \frac{h_T^{1-2\beta}}{k^{3-\beta(4+d)}} \frac{h_T^2}{k^{2(d+2)}} \left(\frac{h_T}{K} \right)^{2(K+1)} \left\{ \|u_T^{(1)}\|^2_{l^1} |u|^2_{K+2,2,T} \right.
\end{aligned}
$$

$$+ \quad |P_{GL}^{K,k}(\nabla I_{GL}^k(u))|_{K+1,2,T}^2 \Bigg\} + C_{12}\frac{C_{13}\max\limits_{x\in T}(\|b\|_{l^2})h_T}{2k}\|P_{GL}'(\nabla e_h^k)\|_{0,2,T}^2.$$

Mit den Definitionen $C_1 = 2C_{12}C_{13}$ und $C_{13} = dC_{11}^2C_{12}C_2^2\max(C_{10}^2, C_{14}^2)\max\limits_{x\in T}(\|b\|_{l^2}), C_{12} > 0$ und der Voraussetzung $\min(1/2, 3/(4+d)) \geq \beta$ folgt weiter

$$\begin{aligned}|\mathrm{III}| \leq & \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\|u_T^{(1)}\|_{l^1}^2\|e_h^k\|_{0,2,T}^2 + \frac{1}{4}\hat{\epsilon}_1\|P_h'(\nabla e_h^k)\|_{0,2,T}^2 \\ & + \frac{h_T^2}{k^{2(d+2)}}\left(\frac{h_T}{K}\right)^{2(K+1)}\|u_T^{(1)}\|_{l^1}^2\left\{|u|_{K+2,2,T}^2 + |P_{GL}^{K,k}(\nabla I_{GL}^k(u))|_{K+1,2,T}^2\right\}.\end{aligned}$$

Aus der Betrachtung des linearen Residuums

$$R(U) \leq \frac{3}{2}\frac{k}{h_T}\max_{\partial^*T\cap R_n}\left(\left|U_T^+ - U_T^-\right|\right) + 2\max(C_0^\Omega, C_0^{\partial\Omega})\frac{k}{h_T}\max_{\partial^*T\setminus R_n}\left(\left|U_T^+ - U_T^-\right|\right) \tag{6.15}$$

ergibt sich insbesondere

$$R(v+w) \leq R(v) + R(w). \tag{6.16}$$

Dies ermöglicht mit entsprechender Wahl von C_{14}, die am Ende der Abschätzung des Terms IV_1 vorgenommen wird, die weitere Behandlung von

$$\begin{aligned}|\mathrm{IV}| \quad \leq \quad & \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\int_T \frac{h_T R(U)^2}{2k^{2d}\max\limits_{x\in T}(\|b\|_{l^2})C_{14}} \\ & \qquad + \frac{C_{14}\max\limits_{x\in T}(\|b\|_{l^2})k^{2d}}{2h_T}C_2^2\left(\frac{h_T}{k^{2+d/2}}\right)^{2(2-\beta)}\|P_h'(\nabla I_{GL}^k(u))\|_{l^2}^2\|P_h'\nabla e_h^k\|_{l^2}^2\,dx \\ \underset{(6.16)}{\leq} & \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\int_T \frac{h_T(R(e_h^k) + R(I_{GL}^k(u)))^2}{2k^{2d}\max\limits_{x\in T}(\|b\|_{l^2})C_{14}} \\ & \qquad + \frac{C_{14}\max\limits_{x\in T}(\|b\|_{l^2})k^{2d}}{2h_T}C_2^2\left(\frac{h_T}{k^{2+d/2}}\right)^{2(2-\beta)}\|P_h'(\nabla I_{GL}^k(u))\|_{l^2}^2\|P_h'(\nabla e_h^k)\|_{l^2}^2\,dx \\ \leq \quad & \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\int_T \frac{h_T R(e_h^k)^2}{k^{2d}\max\limits_{x\in T}(\|b\|_{l^2})C_{14}} + \frac{h_T R(I_{GL}^k(u))^2}{k^{2d}\max\limits_{x\in T}(\|b\|_{l^2})C_{14}} \\ & \qquad + \frac{C_{14}\max\limits_{x\in T}(\|b\|_{l^2})k^{2d}}{2h_T}C_2^2\left(\frac{h_T}{k^{2+d/2}}\right)^{2(2-\beta)}\|P_h'(\nabla I_{GL}^k(u))\|_{l^2}^2\|P_h'(\nabla e_h^k)\|_{l^2}^2\,dx \\ = \quad & \mathrm{IV}_1 + \mathrm{IV}_2 + \mathrm{IV}_3.\end{aligned}$$

$$\begin{aligned}|\mathrm{IV}_1| \quad \leq \quad & \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\frac{h_T^{-1}|T|}{k^{2(d-1)}\max\limits_{x\in T}(\|b\|_{l^2})C_{14}}\frac{9}{4}\max_{\partial^*T\cap R_n}\left(|e_T^+ - e_T^-|\right)^2 \\ + \quad & \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\frac{h_T^{-1}|T|}{k^{2(d-1)}\max\limits_{x\in T}(\|b\|_{l^2})C_{14}}4\left(\max(C_0^\Omega, C_0^{\partial\Omega})\right)^2\max_{\partial^*T\setminus R_n}\left(|e_T^+ - e_T^-|\right)^2\end{aligned}$$

$$
\begin{aligned}
&\underset{(4.61)}{\leq} \sum_{n=0}^{N-1} \frac{9C_{15}^2}{4\max\limits_{x\in T}(\|b\|_{l^2})C_{14}} \|e_+^n - e_-^n\|_{0,2,\Omega}^2 \\
&+ \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \frac{4\left(\max(C_0^\Omega, C_0^{\partial\Omega})\right)^2 C_{15}^2}{\max\limits_{x\in T}(\|b\|_{l^2})C_{14}} \|e_T^+ - e_T^-\|_{0,2,\partial^* T\setminus R_n}^2 \\
&\leq \frac{9C_{15}^2}{4\max\limits_{x\in T}(\|b\|_{l^2})C_{14}} \Big\{ \|e_+^0\|_{0,2,\Omega}^2 + \sum_{n=1}^{N-1} \|e_+^n - e_-^n\|_{0,2,\Omega}^2 \Big\} \\
&+ \frac{4\left(\max(C_0^\Omega, C_0^{\partial\Omega})\right)^2 C_{15}^2}{\max\limits_{x\in T}(\|b\|_{l^2})C_{14}} \sum_{n=0}^{N-1} \Big\{ 2\|e^+ - e^-\|_{0,2,R_{n,n+1}^i}^2 + \|e^+ - e^-\|_{0,2,\Lambda_{n,n+1}}^2 \Big\} \\
&\leq \frac{18C_{15}^2}{8\max\limits_{x\in T}(\|b\|_{l^2})C_{14}} \Big\{ \|e_+^0\|_{0,2,\Omega}^2 + \sum_{n=1}^{N-1} \|e_+^n - e_-^n\|_{0,2,\Omega}^2 \Big\} \\
&+ \frac{64\left(\max(C_0^\Omega, C_0^{\partial\Omega})\right)^2 C_{15}^2}{8\max\limits_{x\in T}(\|b\|_{l^2})C_{14}} \sum_{n=0}^{N-1} \Big\{ \|e^+ - e^-\|_{0,2,R_{n,n+1}^i}^2 + \|e^+\|_{0,2,\Lambda_{n,n+1}^-}^2 \Big\} \\
&\leq \frac{1}{8}\|e_+^0\|_{0,2,\Omega}^2 + \frac{1}{8}\sum_{n=1}^{N-1} \|e_+^n - e_-^n\|_{0,2,\Omega}^2 \\
&+ \frac{1}{8}\sum_{n=0}^{N-1} \Big\{ \|e^+ - e^-\|_{0,2,R_{n,n+1}^i}^2 + \|e^+\|_{0,2,\Lambda_{n,n+1}^-}^2 \Big\},\ C_{15} > 0.
\end{aligned}
$$

Für die analoge Behandlung von IV_2 ist die Ungleichung

$$R(I_{GL}^k(u)) = R(u-\xi) \leq \underbrace{R(u)}_{=0} + R(-\xi) = R(\xi)$$

dienlich und liefert

$$
\begin{aligned}
|\mathrm{IV}_2| &\leq \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \int_T \frac{h_T R(\xi)^2}{k^{2(d-1)}\max\limits_{x\in T}(\|b\|_{l^2})C_{14}}\, dx \\
&\leq \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \frac{1}{8}\|\xi_T^+ - \xi_T^-\|_{0,2,\partial^* T\cap R_n}^2 + \frac{1}{8}\|\xi_T^+ - \xi_T^-\|_{0,2,\partial^* T\setminus R_n}^2 \\
&\leq C\sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} k^\epsilon \frac{h_T^{2k+1}}{k^{2k+1}} |u|_{k+1,2,T}^2, \qquad (6.17) \\
|\mathrm{IV}_3| &\leq \sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \frac{1}{2}\max\limits_{x\in T}(\|b\|_{l^2})C_{14}C_2^2 \frac{h_T^{3-2\beta}k^{2d}}{k^{(2+d/2)(4-2\beta)}} \|P_h'(\nabla I_{GL}^k(u))\|_{0,\infty,T}^2 \|P_h'(\nabla e_h^k)\|_{0,2,T}^2 \\
&\underset{\substack{(6.14),\\(4.61)}}{\leq} C\sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \frac{h_T^{1-2\beta}}{k^{4-\beta(d+4)}} \|u_T^{(1)}\|_{l^1}^2 \|e_h^k\|_{0,2,T}^2 \\
&\leq C\sum_{n=0}^{N-1} \sum_{T\in\mathcal{T}_h^n} \|u_T^{(1)}\|_{l^1}^2 \|e_h^k\|_{0,2,T}^2, \quad \beta \leq \min(\frac{1}{2}, \frac{4}{d+4}).
\end{aligned}
$$

Die Addition aller Abschätzungen ergibt mit $b_1^N = b_1$ für $u \in W^{k+1,2}(Q_T) \cap W^{1,\infty}(Q_T)$

$$
\begin{aligned}
b_1^N(e_h^k, e_h^k) \leq\; & C \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \frac{h_T^2}{k^{2(d+2)}} \left(\frac{h_T}{K}\right)^{2(K+1)} \|u_T^{(1)}\|_{l^1}^2 \left\{ |u|_{K+2,2,T}^2 + |P_{GL}^{K,k}(\nabla I_{GL}^k(u))|_{K+1,2,T}^2 \right\} \\
& + C \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \frac{h_T}{k} \left(\frac{h_T}{K}\right)^{2(K+1)} \left\{ |u|_{K+2,2,T}^2 + |P_{GL}^{K,k}(\nabla I_{GL}^k(u))|_{K+1,2,T}^2 \right\} \\
& + C \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} k^\epsilon \frac{h_T^{2k+1}}{k^{2k+1}} |u|_{k+1,2,T}^2 + C \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \|u_T^{(1)}\|_{l^1}^2 \|e_h^k\|_{0,2,T}^2, \quad 0 < \epsilon \ll 1 \\
= \; & C_{16} \sum_{n=0}^{N-1} \alpha_n + C_{17} \max_{T \in \mathcal{T}_h} \|u_T^{(1)}\|_{l^1}^2 \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \|e_h^k\|_{0,2,T}^2,\; C_{16}, C_{17} > 0,
\end{aligned} \tag{6.18}
$$

falls

$$
0 < \beta \leq \min\left(\frac{1}{2}, \frac{3}{d+4}\right). \tag{6.19}
$$

Eine Möglichkeit die Abhängigkeit der rechten Seite von e_h^k zu entfernen, besteht in der Anwendung von (5.47) für $t_n \leq t \leq t_{n+1}$:

$$
\begin{aligned}
\|e_h^k(t,\cdot)\|_{0,2,\Omega}^2 \;\leq\; & \|e_-^{n+1}\|_{0,2,\Omega}^2 + 2C_1 h \sum_{T \in \mathcal{T}_h^n} \max_{x \in T} \|b\|_{l^2} \int_T \nabla e_h^k \cdot \nabla e_h^k \, dx \\
& + \frac{C_0^\Omega}{2C_1 \min\limits_{T \in \mathcal{T}_h^n}\{h_T\}} \int_t^{t_{n+1}} \|e_h^k(t',\cdot)\|_{0,2,\Omega}^2 \, dt' \\
\leq\; & 2C_6 \left[\sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_1 k \|\nabla e_h^k\|_{0,2,T}^2 + \frac{1}{2} \|e_-^{n+1}\|_{0,2,\Omega}^2 \right] \\
\leq\; & 2C_6 \frac{k}{\tilde{\epsilon}_{\text{vms}}^{\min}} \left[\sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_1 \tilde{\epsilon}_{\text{vms}}^{\min} \|\nabla e_h^k\|_{0,2,T}^2 + \frac{1}{2} \|e_-^{n+1}\|_{0,2,\Omega}^2 \right] \\
\underset{(6.23)}{\leq}\; & 2C_6 \frac{k}{\tilde{\epsilon}_{\text{vms}}^{\min}} \left[\sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_1 \|P_h'(\nabla e_h^k)\|_{0,2,T,h}^2 + \frac{1}{2} \|e_-^{n+1}\|_{0,2,\Omega}^2 \right],
\end{aligned}
$$

so dass weiter folgt

$$
\begin{aligned}
\sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \|e_h^k\|_{0,2,T}^2 &= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \|e_h^k(t,\cdot)\|_{0,2,\Omega}^2 \, dt \\
&\leq \frac{2C_6 k}{\hat{\epsilon}_{\text{vms}}^{\min}} \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \sum_{T \in \mathcal{T}_h^n} \hat{\epsilon}_1 \|P_h'(\nabla e_h^k)\|_{0,2,T,h}^2 + \frac{1}{2} \|e_-^{n+1}\|_{0,2,\Omega}^2 \, dt \\
&\leq \frac{2C_6 k}{\hat{\epsilon}_{\text{vms}}^{\min}} \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} b_1^{n+1}(e_h^k, e_h^k) \, dt \leq \frac{2C_6 k h}{\hat{\epsilon}_{\text{vms}}^{\min}} \sum_{n=0}^{N-1} b_1^{n+1}(e_h^k, e_h^k).
\end{aligned}
$$

Die Anwendung des diskreten Lemmas von Gronwall leistet, falls h hinreichend klein ist

$$2\left(\hat{\epsilon}_{\text{vms}}^{\min}\right)^{-1} C_6 C_{17} k h \max_{T\in\mathcal{T}_h} \|u_T^{(1)}\|_{l^1}^2 \leq C_{18} < 1, \quad C_{18} > 0$$

das Gewünschte:

$$\begin{aligned} b_1^N(e_h^k, e_h^k) &\leq C_{16}\alpha_0 + C_{16}\sum_{n=0}^{N-2}\alpha_{n+1} + C_{18}\sum_{n=0}^{N-1} b_1^{n+1}(e_h^k, e_h^k) \\ &\leq \frac{C_{16}\alpha_0 + C_{16}\sum_{n=0}^{N-2}\alpha_{n+1} + C_{18}\sum_{n=0}^{N-2} b_1^{n+1}(e_h^k, e_h^k)}{1-C_{18}} \\ &\leq \frac{C_{16}\alpha_0 + C_{16}\sum_{n=0}^{N-2}\alpha_{n+1}}{1-C_{18}} \exp\left(\frac{C_{18}(N-1)}{1-C_{18}}\right) \\ &\leq \frac{C_{16}}{1-C_{18}} \exp\left(\frac{N-1}{1-C_{18}}\right)\sum_{n=0}^{N-1}\alpha_n. \end{aligned} \tag{6.20}$$

Aus $|||u-U|||_h \leq \left|\left|\left|e_h^k\right|\right|\right|_h + |||\xi|||_h$ folgt mit der Abschätzung von $|||\xi|||_h$ analog zur Behandlung des Terms I

$$|||\xi|||_h^2 \leq C\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\left(k^{\epsilon}\frac{h_T^{2k+1}}{k^{2k+1}} + \hat{\epsilon}_1\frac{h_T^{2k}}{k^{2k}} + \frac{h_T^{2k+2}}{k^{2k+2}}\right)|u|_{k+1,2,T}^2,$$

die endgültige Fehlerabschätzung

$$|||u-U|||_h^2 \leq C\sum_{n=0}^{N-1}\alpha_n. \tag{6.21}$$

Zusammen mit der

Annahme 6.1.1

1. Es existiert ein $C > 0$ unabhängig von h und k derart, dass gilt
$$|P_{GL}^{K,k}(v)|_{K+1,2,T} \leq C|v|_{K+1,2,T}. \tag{6.22}$$
2. Es gilt die Abschätzung
$$\tilde{\epsilon}_{\text{vms}}^{\min}\|\nabla e_h^k\|_{0,2,T}^2 \leq \|P_h'(\nabla e_h^k)\|_{0,2,T,h}^2 \tag{6.23}$$
mit $\tilde{\epsilon}_{\text{vms}}^{\min} > 0$ unabhängig von h und k.

folgt das

Theorem 6.1.2 *Es sei Ω ein Gebiet mit Lipschitz-Rand und $\{\mathcal{T}_h^n\}_{h>0}$ eine quasiuniforme Familie von Partitionierungen von $(0,T)\times\Omega$. Dann existiert ein $C > 0$ unabhängig von h und k, so dass mit (5.8) und unter der Annahme 6.1.1 für die Lösungen $U \in W_h$ von (5.15) mit Lagrangeschen Finiten-Elementen bzgl. der Gauss-Lobatto Quadraturpunkte und $u \in W^{1,\infty}(Q_T)$ nach Definition 3.2.2 folgende Ungleichung gilt:*

$$|||u-U|||_h \leq C\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\frac{h_T^{1/2}}{k^{1/2}}\frac{h_T^{K+1}}{K^{K+1}}. \tag{6.24}$$

In dem eingangs erwähnten Fall $P_{GL}^{K,k} = \{0\}$ ist die Konvergenz des Verfahrens ebenso sichergestellt.

Theorem 6.1.3 *Es sei Ω ein Gebiet mit Lipschitz-Rand und $\{\mathcal{T}_h^n\}_{h>0}$ eine quasiuniforme Familie von Partitionierungen von $(0,T)\times\Omega$. Dann existiert ein $C>0$ unabhängig von h und k, so dass mit (5.8) für die Lösungen $U\in W_h$ von (5.15) mit $P_{GL}^{K,k}=\{0\}$ und Lagrangeschen Finiten-Elementen bzgl. der Gauss-Lobatto Quadraturpunkte folgende Ungleichung gilt:*

$$|||u-U|||_h \le C\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\frac{h_T^{1/2}}{k^{1/2}}. \tag{6.25}$$

$u\in W^{1,\infty}(Q_T)$ ist hierbei eine Lösung nach Definition 3.2.2.

Beweis Analog zu (6.21). □

6.2 A priori Fehlerabschätzung II

Gegeben sei die schwache Formulierung (vgl. [JJS95, (2.7)]):

Finde $U\in W_h$ derart, dass für $n=0,1,\dots,N-1$, $U\equiv U_{Q_{n,n+1}}\in W_h^n$

$$a(U,v)+\sum_{T\in\mathcal{T}_h^n}\left\{\delta\left(L(U),f'(U)\cdot\nabla v\right)_{0,T}+\hat{\epsilon}\left(\nabla U,\nabla v\right)_{0,T}\right\}=0\quad\forall v\in W_h^n \tag{6.26}$$

mit $L(U)=\nabla\cdot f(U)$, $0<\beta<\min\left(\frac{1}{2},\frac{4}{d+4}\right)$ und

$$\delta=\delta(U)=C_1\frac{h_T}{k}\left(\max_{x\in T}\|f'(U)\|_{l^2}\right)^{-1},$$
$$\hat{\epsilon}=\hat{\epsilon}(U)=\max\left(C_2\left(\frac{h}{k^{2+d/2}}\right)^{2-\beta}R(U),C_3\left(\frac{h}{k}\right)^{k+1/2}\right),$$
$$R(U)|_T=\max_T\left(|L(U)|\right)+\frac{k}{h_T}\left(\max_{\partial^*T}\left(\left|\left[f(U_T^+)-f(U_T^-)\right]\cdot n_T^+\right|\right)+\max_{\partial^*T}\left(C_T\left|U_T^+-U_T^-\right|\right)\right).$$

Nahezu analog zum vorherigen Abschnitt kann eine weitere a priori Fehlerabschätzung für das lineare Problem mit (6.5) und der Annahme (6.6) gezeigt werden. Ausgangspunkt sind die Gleichungen

$$b_1(u,v)=l(v)\quad\forall v\in W^{1,2}(Q_T,\mathcal{T}_h) \tag{6.27}$$

und

$$b_1(u,v)=l(v)\quad\forall v\in W_h, \tag{6.28}$$

mit

$$
\begin{aligned}
b_1(v,w) = &\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\int_T \nabla\cdot f(v)w\,dx + \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\delta(v)\left(L(v), b\cdot\nabla w\right)_{0,T}\\
&+\sum_{n=1}^{N-1}\int_\Omega (v_+^n - v_-^n)w_+^n\,dx_x + \int_\Omega v_+^0 w_+^0\,dx_x\\
&+\sum_{n=0}^{N-1}\int_{R_{n,n+1}^i}\frac{1}{2}\left(f(v^-)-f(v^+)\right)\cdot n^+(w^+ + w^-)\,ds\\
&+\sum_{n=0}^{N-1}\int_{R_{n,n+1}^i} C_0^\Omega (v^+ - v^-)(w^+ - w^-)\,ds\\
&-\sum_{n=0}^{N-1}\int_{\Lambda_{n,n+1}^-}\frac{1}{2}f(v^+)\cdot n^+ w^+\,ds + \sum_{n=0}^{N-1}\int_{\Lambda_{n,n+1}^-} C_0^{\partial\Omega} v^+ w^+\,ds.
\end{aligned}
\tag{6.29}
$$

Weitere Zwischenergebnisse auf dem Weg zum nächsten Theorem sind

$$
b_1(e_h^k, e_h^k) = b_1(\xi, e_h^k) - \sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\hat{\epsilon}(U)\left(\nabla I_{GL}^k(u), \nabla e_h^k\right)_{0,T} = \mathrm{I} + \mathrm{II},
$$

$$
|\mathrm{I}| \le \frac{1}{4}b_1(e_h^k, e_h^k) + C\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n} k^\epsilon \frac{h_T^{2k+1}}{k^{2k+1}}|u|_{k+1,2,T}^2, \quad 0<\epsilon\ll 1,
$$

$$
\begin{aligned}
|\mathrm{II}| \le &\frac{1}{8}\|e_+^0\|_{0,2,\Omega}^2 + \frac{1}{8}\sum_{n=1}^{N-1}\|e_+^n - e_-^n\|_{0,2,\Omega}^2 + \frac{1}{8}\delta\|L(e_h^k)\|_{0,2,T}^2\\
&+\frac{1}{8}\sum_{n=0}^{N-1}\left\{\|e^+ - e^-\|_{0,2,R_{n,n+1}^i}^2 + \|e^+\|_{0,2,\Lambda_{n,n+1}^-}^2\right\}\\
&+C\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n} k^\epsilon\frac{h_T^{2k+1}}{k^{2k+1}}|u|_{k+1,2,T}^2\\
&+C\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n}\|u_T^{(1)}\|_{l^1}^2\|e_h^k\|_{0,2,T}^2, \quad \beta\le\min(\frac{1}{2},\frac{4}{d+4}).
\end{aligned}
$$

Theorem 6.2.1 *Es sei Ω ein Gebiet mit Lipschitz-Rand und $\{\mathcal{T}_h^n\}_{h>0}$ eine quasiuniforme Familie von Partitionierungen von $(0,T)\times\Omega$. Dann existiert ein $C>0$ unabhängig von h und k, so dass für die Lösungen $U\in W_h$ von (6.26) mit Lagrangeschen Finiten-Elementen bzgl. der Gauss-Lobatto Quadraturpunkte und $u\in W^{1,\infty}(Q_T)$ nach Definition 3.2.2 folgende Ungleichung gilt:*

$$
|||u-U|||_h \le C\sum_{n=0}^{N-1}\sum_{T\in\mathcal{T}_h^n} k^\epsilon\frac{h_T^{k+1/2}}{k^{k+1/2}},\ 0<\epsilon\ll 1. \tag{6.30}
$$

7 Numerische Beispiele

In diesem Kapitel werden zwei Beispiele mit den Verfahren (5.15) numerisch gelöst, um zum einen die erreichte Konvergenzordnung mit der Aussage des Theorems $6.1.2$ zu vergleichen und zum anderen die Qualität der stabilisierten Lösung zu beurteilen. Für diese numerische Untersuchungen werden für das stationäre Modellproblem mit externen Quellterm f

$$L(u) = b \cdot \nabla u = f \text{ in } \Omega, \tag{7.1}$$

$$u = g_D \text{ auf } \Gamma_D \tag{7.2}$$

Beispiele mit und ohne Grenzschichten verwendet.

Die Diskretisierung von (5.15) benutzt die Parameter $C_0^{\partial\Omega} = C_0^{\Omega} = 3.0$, $C_1 = C_2 = C_3 = 0.1$ und $\beta = 0.1$ aus (5.19). Für Vergleichszwecke wird die Methode (5.34) mit einer vergleichbaren Parameterwahl $C_0^{\partial\Omega} = C_0^{\Omega} = 3.0$, $C_1, C_2 = 0.1, C_3 = 0$ und $\beta = 0.1$ ebenfalls zur Lösung der Beispiele herangezogen.

Bei den numerischen Tests wird auch der Fluktuationsoperator $P_h' = I - P_h^{0,k}$ berücksichtigt, obwohl er nicht in die Theorie von Kapitel 4 eingegliedert ist, da die Gauss-Lobatto-Quadraturpunkte erst ab $1 \leq k$ definiert sind (vgl. (4.42)). Dessen ungeachtet bilden die Lagrange-Polynome vom Grad 2 eine eingebettete hierarchische nodale Basis vom Grad 0.

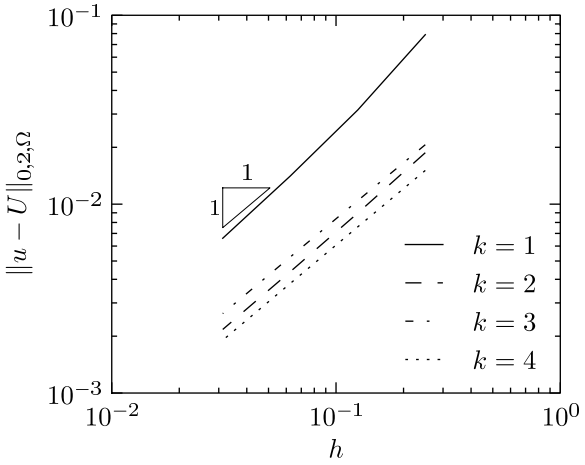
7.1 Numerische Konvergenzuntersuchung

Beispiel 7.1.1 (Glatte Lösung ohne Grenzschichten) (Vgl. [HJS02, Example 4.2]) Im ersten Beipiel ist $\Omega = (-1, 1)^2$, $(b_1(x, y), b_2(x, y)) = (8/10, 6/10)$ und $g_D = 1$. Die rechte Seite wird so gewählt, das die analytische Lösung

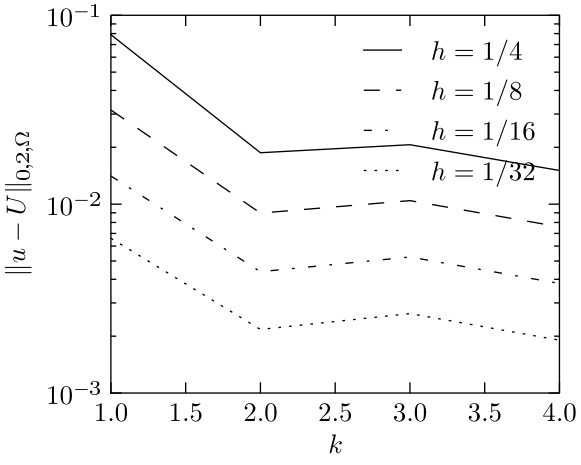
$$u(x, y) = 1 + \sin(\pi(1 + x)(1 + y)^2/8)$$

ist.

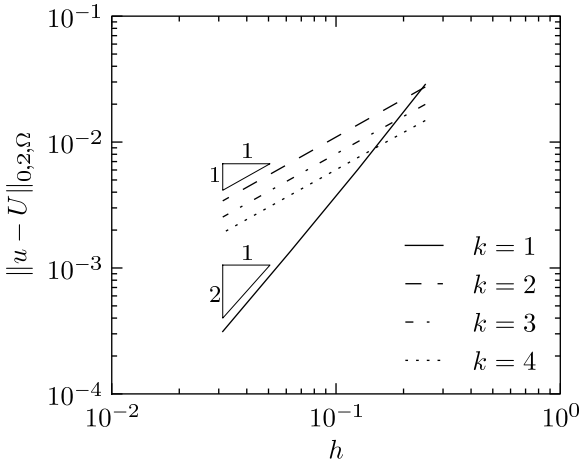
Für die exemplarische Konvergenzuntersuchung am Beispiel 7.1.1 werden die Gitterweiten $h \in \{1/4, 1/8, 1/16, 1/32\}$ und die Polynomgrade $k \in \{1, 2, 3, 4\}$ gewählt. In den Abbildungen 7.1a und 7.1c wird deutlich, dass der Fehler $\|u - U\|_{0,2,\Omega}$ der Methode (5.15) für jedes betrachtete k mit $\mathcal{O}(h)$ für $h \to 0$ konvergiert. Die Ausnahme für $k = 1$ in der Abbildung 7.1c liefert einen Fehler der Form $\mathcal{O}(h^2)$. Dies ist jedoch der Fehler der unstabilisierten Discontinuous-Galerkin-Methode, denn mit $K = k = 1$ gilt $P_h' = I - P_h^{1,1} = 0$.



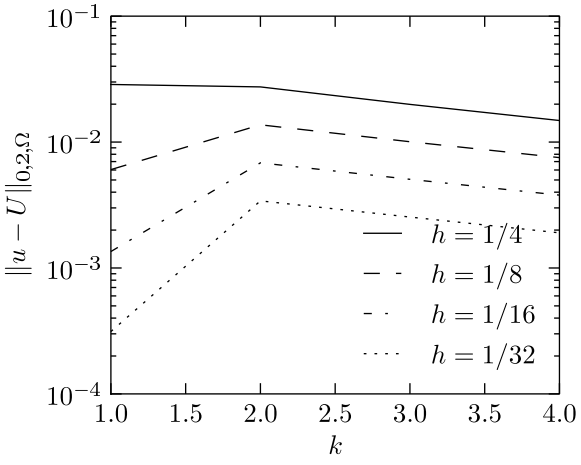
(a) Konvergenz von (5.15), $P'_h = I - P_h^{0,k}$ mit h-Verfeinerung.



(b) Konvergenz von (5.15), $P'_h = I - P_h^{0,k}$ mit k-Verfeinerung.



(c) Konvergenz von (5.15), $P'_h = I - P_h^{1,k}$ mit h-Verfeinerung.



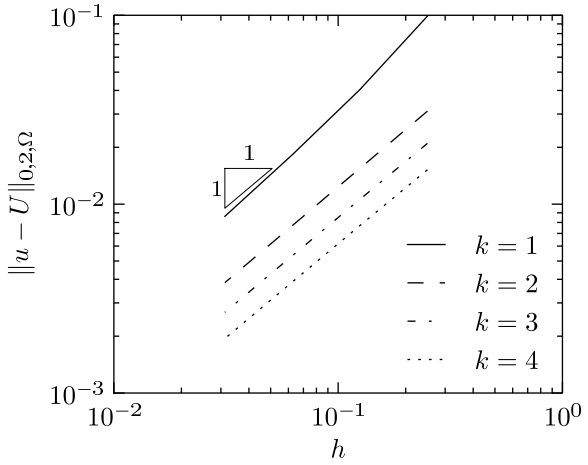
(d) Konvergenz von (5.15), $P'_h = I - P_h^{1,k}$ mit k-Verfeinerung.

Abbildung 7.1: Beispiel 1

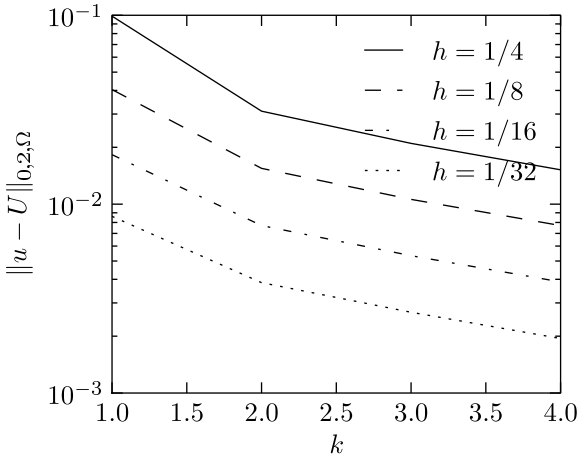
Darüber hinaus wird in den Abbildungen 7.1b und 7.1d bei fixierter Gitterweite h die Abhängigkeit des Fehler vom Polynomgrad k betrachtet. Da sich in den numerischen Ergebnissen keinerlei Übereinstimmung mit den Aussagen des Theorems 6.1.2 entdecken lässt, ist davon auszugehen, dass die Annahme 6.1.1 nicht befriedigt werden kann.

Die Abbildungen 7.2a und 7.2b behandeln die Ergebnisse für den Fall, dass der Shock-capturing Term auf alle Skalen wirkt und gehen konform mit dem Theorem 6.1.3.

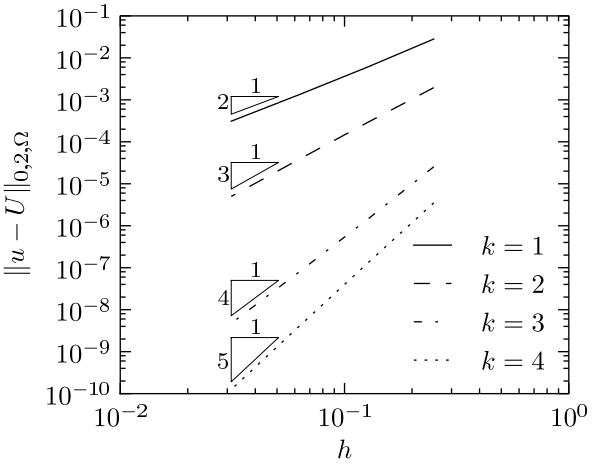
Als Abschluss der numerischen Konvergenzuntersuchung liefern die letzten beiden Abbildungen die Ergebnisse des Verfahrens (6.26). In ihnen findet sich das theoretische Resultat des Theorems 6.2.1 wieder. Abbildung 7.2d liefert, unter Berücksichtigung der Skalierung, mit den angenäherten Geraden des Konvergenzplots einen Indiz für exponentielle Konvergenz der diskreten Lösung bzgl. des Polynomgrades.



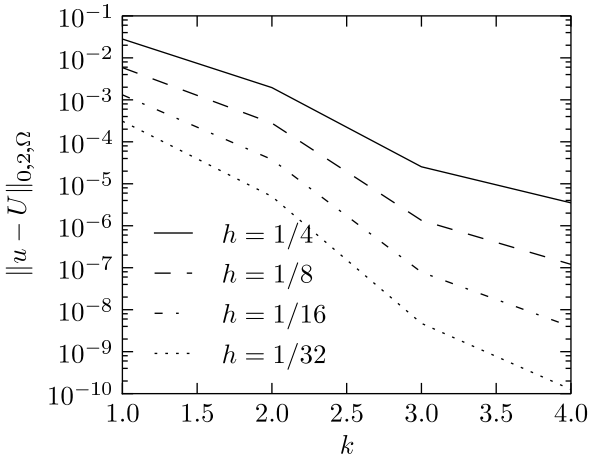
(a) Konvergenz von (5.15), $P_h' = I$ mit h-Verfeinerung.



(b) Konvergenz von (5.15), $P_h' = I$ mit k-Verfeinerung.



(c) Konvergenz von (6.26) mit h-Verfeinerung.



(d) Konvergenz von (6.26) mit k-Verfeinerung.

Abbildung 7.2: Beispiel 1

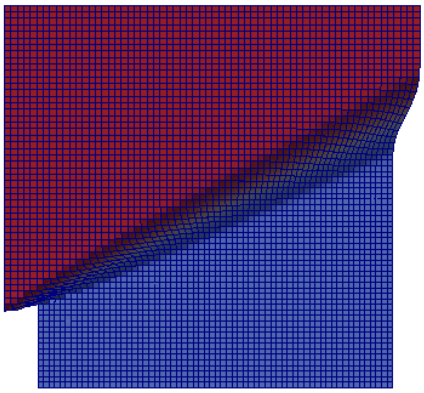
7.2 Modellproblem mit Grenzschichten

Beispiel 7.2.1 (Innere Grenzschicht) (Vgl. [Ang95, Problem 1]) In diesem Fall ist $\Omega = (0,1)^2$, $(b_1(x,y), b_2(x,y)) = (2,1)$ und $f = 0$. Die Randbedingung genügt der folgenden Funktion

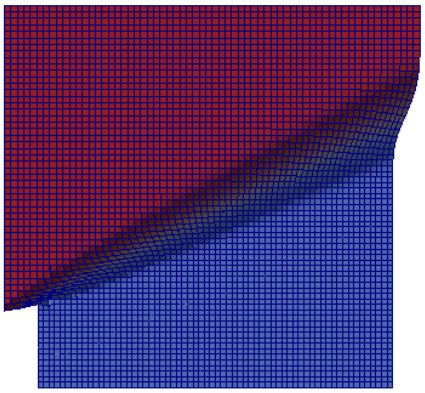
$$g(x) = \begin{cases} 0, & x_2 < \frac{3}{4}x_1 + \frac{1}{4} \\ \frac{1}{2}, & x = (0, \frac{1}{4}) \\ 1, & \text{sonst.} \end{cases}$$

Die numerischen Lösungen sind in Abbildungen 7.3a-7.3d dargestellt. Im Vergleich der beiden Lösungen mit dem Fluktuationsoperatoren $P_h' = I - P_h^{0,2}$ und $P_h' = I - P_h^{1,2}$ fällt auf, dass die erste Lösung eine steilere Grenzschicht im Ausströmrand besitzt. Dieses Verhalten erscheint insofern merkwürdig, als dass $I - P_h^{0,2}$ mehr Skalen bei der Stabilisierung erfasst als

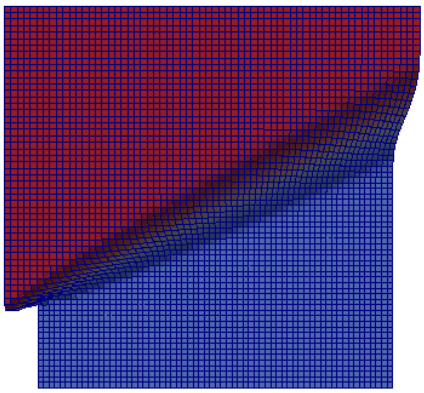
$I - P_h^{1,2}$. Bei der Berücksichtigung von sämtlichen Skalen mit $P_h' = I$ in Abbildung 7.3c, entsteht jedoch wieder eine Grenzschicht von der Form aus Abbildung 7.3b. Die Grenzschicht im Einströmrand ist hier dagegen etwas breiter als in den Abbildungen 7.3a und 7.3b. Der abschließende Vergleich mit der Methode (6.26) zeigt, dass die entsprechende Lösung eine deutlich steilere Grenzschicht besitzt als die Methoden ohne Stromliniendiffusionsterm, aber im Gegensatz dazu auch mit Oszillation aufwartet.



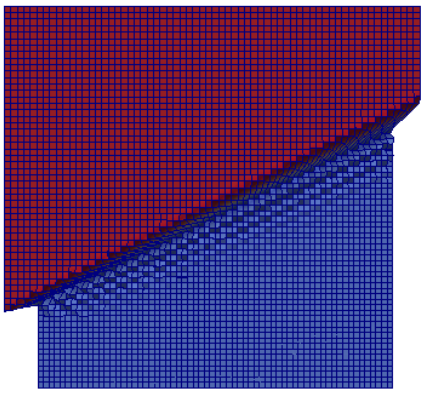
(a) Methode (5.15) mit $P_h' = I - P_h^{0,2}$.



(b) Methode (5.15) mit $P_h' = I - P_h^{1,2}$.



(c) Methode (5.15) mit $P_h' = I$.



(d) Methode (6.26).

Abbildung 7.3: Beispiel 2

8 Zusammenfassende Diskussion und Ausblick

In dieser Arbeit wurden Discontinuous-Galerkin-Methoden zur Lösung von hyperbolischen Problemen erster Ordnung betrachtet, die zur Vermeidung von nichtphysikalischen Oszillationen verschiedene Arten von künstlicher Diffusion einsetzen.

Ziel war es, eine $L^\infty(L^\infty)$-Abschätzung der mit diesen Verfahren gewonnenen diskreten Lösung zu erhalten. Mit dem Lemma 5.4.8 wurde in Verbindung mit den Ergebnissen aus [JSH90], [Sze89a] oder [Sze91] diese Abschätzung für die Lösung der Methode (6.26) bzw. [JJS95, (2.7)] mit Lagrangeschen Finiten-Elementen höherer Ordnung auf quasiuniforme Partitionierungen bewiesen. Bisher gelang für $d_x = 1$ in [JSH90] und [Sze91] der Nachweis der entsprechenden Ungleichung nur in einer schwächeren Form für lineare Funktionen. Für dieses Ergebnis existiert in [Sze89a] lediglich eine Erweiterung für Dreiecke mit einem rechten Winkel $(d_x = 2)$, jedoch ebenfalls nur für $k = 1$. Diese Beschränkung wird in [Sze91] umgangen, indem der Shock-capturing Term auf eine feinere Triangulierung mit linearen Ansatzfunktionen interpoliert wird, so dass die $L^\infty(L^\infty)$-Abschätzung zumindest für diese Modifikation garantiert werden kann (vgl. Bemerkung 5.4.13).

Ferner wurde ein Verfahren konstruiert, das im Gegensatz zu (6.26) keine Diffusion in Stromlinienrichtung benötigt, um eine $L^\infty(L^\infty)$-Abschätzung zu garantierten. Die dafür vorgenommene Modifikation an der nichtlinearen, isotropen Diffusion zerstört jedoch den residualen Charakter der Methode. Um eine Konvergenzordnung, die größer als Eins ist, zu erhalten, wurde die Idee verfolgt, die künstliche Diffusion nur auf feine Skalen wirken zu lassen (vgl. [Hei07]). Für die $L^\infty(L^\infty)$-Abschätzung, ist an dem dazu notwendigen Fluktuationsoperator P_h' die Bedingung

$$\left(P_h'(\nabla v), P_h'(\nabla v^{p-1})\right)_{0,T,h} \geq 0 \tag{8.1}$$

zu stellen. Der aus dieser Bedingung hervorgegangene diskrete Projektor P_h', besitzt durch die Eigenschaften der Gauss-Lobatto-Interpolation auf den ersten Blick geeignete Fehlerdarstellungen (Korollar 4.4.8) - so lange nur die Stetigkeit des grobskaligen Projektors $P_h^{K,k}$ bzgl. entsprechender Halbnormen erfüllt ist (Annahme 6.1.1 (1)).

Eine weitere Problematik findet sich im zweiten Teil der Annahme 6.1.1:

$$\tilde{\epsilon}_{\text{vms}}^{\min} \|\nabla e_h^k\|_{0,2,T}^2 \leq \|P_h'(\nabla e_h^k)\|_{0,2,T,h}^2.$$

Falls der Projektor P_h' dieses koerzive Verhalten aufweisen sollte, so ist doch die Unabhängigkeit der Konstante von h und k zumindest fraglich. Erforderlich ist dieser Teil der Annahme, da die Konvergenzanalysis die Ungleichung

$$b_1^N(e_h^k, e_h^k) \leq \cdots + C \sum_{n=0}^{N-1} \sum_{T \in \mathcal{T}_h^n} \|e_h^k\|_{0,2,T}^2$$

aufweist, die ihren Ursprung in der Nichtlinearität des Shock-capturing Terms hat. Ein Vergleich mit [Hei07], bei der die feinskalige Projektion zu akzeptablen a priori Fehlerdarstellungen führt, zeigt: die Ursachen beider Annahmen

- $L^\infty(L^\infty)$-Abschätzung
- Nichtlinearität des Konvektionsterms

treten dort nicht auf.

Da die numerischen Vergleichsrechnungen am Beispiel $k = 2$, $K = 0,1$ nur einen Fehler der Ordnung $\mathcal{O}(h)$ aufweisen, ist zu vermuten, dass wenigstens eine der beiden Annahmen nicht erfüllt ist. Sollte die Verletzung der zweiten Annahme die Ursache für die geringe Konvergenzordnung sein, so kann die Projektion im nichtlinearen Shock-capturing Term kein zufriedenstellendes Verfahren liefern. Andernfalls könnte eine Änderung des Projektors, unter Berücksichtigung von (8.1), der vorgestellten Methode zum Erfolg verhelfen.

Beide Annahmen könnten umgangen werden, wenn es gelänge die isotrope Diffusion so zu modifizieren, dass eine $L^\infty(L^\infty)$-Abschätzung unter Beibehaltung der residualen Struktur möglich ist. Die Motivation zur Einführung einer lokalen Projektion wäre dann nicht mehr gegeben.

Bisher unbeantwortet geblieben ist auch die Frage nach der Konvergenz der Discontinuous-Galerkin-Methode für nichtlineares f gegen eine Lösung von $(3.12)-(3.14)$. Diese Konvergenz kann mit Hilfe der $L^\infty(L^\infty)$-Beschränkung sichergestellt werden, falls eine *maßwertige Lösung* nach Definition [Sze89b, S. 184] existiert.

Index

Symbolverzeichnis

Basisnotationen

$\emptyset$	leere Menge
$\mathbb{N}_0$	natürliche Zahlen $\{0, 1, 2, \ldots\}$
$\mathbb{N}$	positive natürliche Zahlen $\{1, 2, \ldots\}$
$\mathbb{Z}$	ganze Zahlen $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$
$\mathbb{R}$	reelle Zahlen $(-\infty, +\infty)$
$\mathbb{R}_+$	positive reelle Zahlen $(0, +\infty)$
$\#\mathcal{M}$	Kardinalzahl einer Menge $\mathcal{M}$
$\dim(V)$	Dimension des Vektorraums V
$v\|_G$	Restriktion der Funktion v auf die Menge G
$\|G\|$	Lebesgue-Maß von $G \subset \mathbb{R}^n$, $n \in \mathbb{N}$
∂G	Rand von G
$\operatorname{span}\{v_1, \ldots, v_n\}$	Lineare Hülle der Vektoren $v_1, \ldots, v_n$, $n \in \mathbb{N}$
$\operatorname{conv}\{v_1, \ldots, v_n\}$	Konvexe Hülle der Vektoren $v_1, \ldots, v_n$, $n \in \mathbb{N}$
δ_{ij}	Kronecker-Symbol: $\delta_{ij} = 1$ falls $i = j$ und 0 sonst
d_x	$\in \mathbb{N}$ Dimension der räumlichen Grundmenge
d	$= d_x + 1$
$\lfloor l \rfloor$	$= \max_{z \in \mathbb{Z},\, z \le l}(z)$
$\lceil l \rceil$	$= \min_{z \in \mathbb{Z},\, z \ge l}(z)$
Ω	$\subset \mathbb{R}^{d_x}$ ein beschränktes Gebiet mit Lipschitz-Rand
Q_T	$= (0, T) \times \Omega \subset \mathbb{R}_+ \times \mathbb{R}^{d_x}, T > 0$ ein Zeit-Raum-Zylinder, S. 19
Σ_T	$= (0, T) \times \partial\Omega, T > 0$ Rand des Zeit-Raum-Zylinders, S. 19
Γ_D	$= \partial\Omega$ Dirchlet-Rand
Γ_D^-	Einflussrand, S. 19
Γ_D^+	$= \Gamma_D \setminus \Gamma_D^-$ Ausflussrand, S. 19
Σ_D^-	$= (0, T) \times \Gamma_D^-, T > 0$, S. 19
Σ_D^+	$= (0, T) \times \Gamma_D^+, T > 0$, S. 19
n	äußerer Normaleneinheitsvektor
C	> 0, allgemeine Konstante, die nach Bedarf angepasst wird
$\tilde{C}, C_0, \cdots$	Konstanten
$\mathcal{L}_{[f]}$	Lipschitz-Konstante der Funktion f bzgl. einer Kugel, S. 25
$\operatorname{dist}(A, B)$	$= \inf_{x \in A, y \in B} \|x - y\|$
$\mathcal{I}[a, b]$	$= [\min(a, b), \max(a, b)]$, S. 23

Vektoren und Matrizen

$(x_1, \ldots, x_n)^T$	Komponenten eines Vektors $x \in \mathbb{R}^n, n \in \mathbb{N}$
e_i	Einheitsvektor $1 \leq i \leq d$
α	$= (\alpha_1, \ldots, \alpha_n)^T \in \mathbb{N}_0^n$ Multiindex
$\lvert\alpha\rvert$	$= \sum_{i=1}^n \alpha_i$ Länge des Multiindex
$\alpha!$	$= \alpha_1! \cdots \alpha_m!, m \in \mathbb{N}$
x^α	$= x_1^{\alpha_1} \cdots x_m^{\alpha_m}$, $m \in \mathbb{N}$
$\mathcal{I}$	Indexmenge
$\lVert v\rVert_{l^p(\mathcal{I})}$	$= (\sum_{i\in\mathcal{I}} \lvert v_i\rvert^p)^{1/p}, p \in [1, \infty)$, $v = (v_i)_{i\in\mathcal{I}}$ Vektorraumnorm
$\lVert v\rVert_{l^\infty(\mathcal{I})}$	$= \max\limits_{i\in\mathcal{I}}\{\lvert v_i\rvert\}$, $v = (v_i)_{i\in\mathcal{I}}$ vektorielle Maximumsnorm
$(v, w)_{l^2(\mathcal{I})}$	$= \sum_{i\in\mathcal{I}} v_i w_i$, $v = (v_i)_{i\in\mathcal{I}}$, $w = (w_i)_{i\in\mathcal{I}}$
$\lVert v\rVert_{l^p}, (v, w)_{l^2}$	$\lVert v\rVert_{l^p(\mathcal{I})}$ bzw. $(v, w)_{l^2(\mathcal{I})}$, $p \in [1, \infty]$
	mit $\mathcal{I} = \{0, 1, 2, \ldots, n-1\}$ oder $\mathcal{I} = \{1, 2, \ldots, n\}$, $n = \#\mathcal{I}$
$\lVert f\rVert_{l^p(\mathcal{M})}$	$= (\sum_{x\in\mathcal{M}} \lvert f(x)\rvert^p)^{1/p}, p \in [1, \infty)$ diskrete Norm
	für Funktionen auf einer Punktmenge $\mathcal{M}$
$x \parallel y_f$	duale Vektoren x, y_f; sie erfüllen für $1/p + 1/q = 1$ die Bedingung
	$1 = x^T y_f = \lVert x\rVert_{l^p} \lVert y_f\rVert_{l^q}$, S. 67
A	$= (a_{ij})_{ij} \in \mathbb{R}^{m,n}$, $m, n \in \mathbb{N}$ Matrix
I	Einheitsmatrix
A^T	Transponierte der Matrix A
$\mathrm{diag}(v)$	Diagonalmatrix, $v \in \mathbb{R}^n, n \in \mathbb{N}$
$\det$	Determinatenfunktion
$\lVert A\rVert_{l^p}$	von $\lVert x\rVert_{l^p}$, $x \in \mathbb{R}^n$, $1 \leq p \leq \infty$ induzierte Matrixnorm
$\lVert A\rVert_{\max}$	$= \max\limits_{1\leq i,j\leq n} \lvert a_{ij}\rvert$
$\mathrm{rg}(A)$	Rang der Matrix A
$\sigma(A)$	Spektrum der Matrix A
$\lambda_{\min}, \lambda_{\max}, \lambda_i$	Eigenwerte einer entsprechenden Matrix
$W(A, \lVert\cdot\rVert)$	numerischer Wertebereich der Matrix A bzgl. der Norm $\lVert\cdot\rVert$, S. 67
$\mathcal{A}$	normierte Algebra mit Einselement, S. 68
$V_{\mathcal{A}}(a, \lVert\cdot\rVert)$	algebraischer numerischer Wertebereich von $a \in \mathcal{A}$, S. 68

Operatoren

$\partial_t u$	$= \frac{\partial u}{\partial t}$ Zeitableitung von u
$\partial_i u$	$= \frac{\partial u}{\partial x_i}$ partielle Ableitung nach x_i, $i \in \mathbb{N}_0$ von u
$\partial_\eta u$	partielle Ableitung von u in Richtung η
$\partial_{ij} u$	zweite Ableitung von u bzgl. x_i und x_j
$\partial^\alpha u$	$= \partial_1^{\alpha_1} \cdots \partial_m^{\alpha_m}$, $m \in \mathbb{N}$
∇u	Gradient von u
$\nabla \cdot u$, $\operatorname{div} u$	Divergenz von u
Δu	Laplace-Operator
L	Differentialoperator
η, q	Entropie-Flux Paar, S. 23
$\operatorname{sign}(\cdot)$	Vorzeichenfunktion
$\operatorname{sign}^+(\cdot)$	$= 1$ falls Argument positiv, sonst 0
$\operatorname{sign}^-(\cdot)$	$= -1$ falls Argument negativ, sonst 0
E	Erweiterungsoperator, S. 12
$\tilde{\gamma}, \gamma, \gamma_1, \gamma_2$	Spuroperatoren, S. 11, 13, 20
L_H	Lumpingoperator, S. 37
I_h^k	Lagrangescher Interpolationsoperator
I_G^k	I_h^k bzgl. der Gauss-Quadraturpunkte
I_{GL}^k	I_h^k bzgl. der Gauss-Lobatto-Quadraturpunkte
P_h^k	L^2-Projektion
$P_h^{K,k}, P_{GL}^{K,k}$	Projektionen auf einem abstrakten Grobraum, S. 42
P_h'	$= I - P_h^{K,k}$ Fluktuationsoperator, S. 42
H	Flussdichte, S. 51
R	Residualoperator bzgl. der Kanten der künstlichen Diffusion, S. 55

Funktionenräume

l	$\in \mathbb{N}_0$ Glattheitsindex für Funktionenräume
k	$\in \mathbb{N}_0$ Polynomgrad
P	allgemeiner Polynomraum
$S(X)$	Einheitssphäre des normierten Raumes X
X'	Dualraum von X
$\mathbb{Q}_k$	$= \operatorname{span}\{x^\alpha : \alpha \in \mathbb{N}_0^d,\ \lvert\alpha\rvert \leq k\},\ k \in \mathbb{N}_0$
$C^l(G)$	$= C^l(G, \mathbb{R}), l \in \mathbb{N}_0$ Raum der l-mal stetig differenzierbaren Funktionen auf $G \subset \mathbb{R}^m, m \in \mathbb{N}$ in die reellen Zahlen
$C_0^l(G)$	Funktionen aus $C^l(G), l \in \mathbb{N}_0$ deren Support kompakt ist
$L^p(G)$	$= L^p(G, \mathbb{R}), p \in [1, \infty)$ Funktionen von G in die reellen Zahlen deren p-te Potenz Lebesgue-integrierbar ist
$L^\infty(G)$	$= L^\infty(G, \mathbb{R})$ Raum der wesentlich beschränkten Funktionen
$W^{l,p}(G)$	$= W^{l,p}(G, \mathbb{R})$ Funktionen deren l-te Ableitungen in $L^p(G)$ liegen
$BV(G)$	$= BV(G, \mathbb{R}) \subset L^1(G, \mathbb{R})$ Funktionen mit beschränkter Variation
$\mathcal{L}(A, B)$	Raum der stetigen, linearen Funktionen von A nach B
$\lVert u\rVert_{L^p(G)}$	$= (\int_G \lvert u\rvert^p\, dx)^{1/p}, p \in [1, \infty)$
$\lVert u\rVert_{L^\infty(G)}$	$= \inf\{K > 0 : \lvert u(x)\rvert \leq K$ fast überall in $G\}$
$\lVert u\rVert_{l,p,G}$	$= (\sum_{\lvert\alpha\rvert \leq l} \lVert \partial^\alpha u\rVert^p_{L^p(G)})^{1/p}, p \in [1, \infty)$
$(u, v)_{l,G}$	$= \sum_{\lvert\alpha\rvert \leq l} \int_G \partial^\alpha u\, \partial^\alpha v\, dx$
$\lVert u\rVert_{l,\infty,G}$	$= \max\{\lVert \partial^\alpha u\rVert_{L^\infty(G)} : \lvert\alpha\rvert \leq l\}$
$\int_G \lvert \mathcal{G} f\rvert dx$	$= \sup\{\int_G f\, div\, g\, dx : g = (g_1, \ldots, g_n)^T \in C_0^1(G)^n,\ \lvert g(x)\rvert \leq 1,\ x \in G\} < \infty$

Symbole im Finite Elemente Kontext

$Q_{n,n+1}$	$= (t_n, t_{n+1}) \times \Omega,\ n \in \mathbb{N}_0$, S. 27
Q_n	$= \{t_n\} \times \Omega,\ n \in \mathbb{N}_0$, S. 27
$\Sigma^{\pm}_{n,n+1}$	$= (t_n, t_{n+1}) \times \Gamma^{\pm}_D,\ n \in \mathbb{N}_0$, S. 27
$\Sigma^{\pm}_n$	$= \{t_n\} \times \Gamma^{\pm}_D,\ n \in \mathbb{N}_0$, S. 27
$R_{n,n+1}, R_n$	Kanten in $Q_{n,n+1}$, Q_n, S. 27
$\Lambda^{\pm}_{n,n+1}, \Lambda^{\pm}_n$	Kanten in $\Sigma^{\pm}_{n,n+1}$, $\Sigma^{\pm}_n$, S. 27
$R^i_{n,n+1}$	$= R_{n,n+1} \setminus \Lambda_{n,n+1}$, S. 27
$\mathcal{T}^n_h$	Zeit-Raum Triangulierung von $Q_{n,n+1}$
h_T	$= \operatorname{diam}(T) = \max_{x,y \in T} \|x - y\|_{l^2}$, Durchmesser von T
ρ_T	Durchmesser der größten in T enthaltenen Kugel
h	$= \max\{h_T : T \in \mathcal{T}^n_h,\ n = 1, \ldots, N,\ N \in \mathbb{N}\}$
$\{\hat{T}, \hat{P}, \hat{\Sigma}\}$	Finites Referenzelement
$\{T, P, \Sigma\}$	Finites Element
$n^k_{\text{dof}}, n^K_{\text{dof}}$	$= \#\hat{\Sigma}$ Freiheitsgrade des Finiten Referenzelementes
$\mathcal{N}$	Knotenmenge eines Lagrangeschen Finiten Elements
$\mathcal{Q}$	Knotenmenge einer nodalen Quadraturformel
$(u, v)_{l,T,h}$	$= \sum_{\|\alpha\| \le l} \sum_{i=1}^{n_{dof}} \partial^\alpha u(x_i)\, \partial^\alpha v(x_i) \omega^J_i,\ x_i \in \mathcal{Q}$
$(u, v)_{l,T,GL}$	$= (u, v)_{l,T,h}$, falls $\mathcal{Q}$ aus den Gauss-Lobatto-Quadraturpunkten besteht
$\|u\|_{l,p,T,h},\ \|u\|_{l,p,T,GL}$	diskrete Normen bzgl. entsprechender Quadraturpunkte, S. 37
W^n_h	$= \{w \in L^2(Q_{n,n+1}) : w\|_T \in P(T)\, \forall T \in \mathcal{T}^n_h\}$
W_h	$= \prod_{n \ge 0} W^n_h$
$W^{l,p}(Q_T, \mathcal{T}_h)$	$\prod_{n \ge 0} \left\{w \in L^2(Q_{n,n+1}) : w\|_T \in W^{l,p}(T)\, \forall T \in \mathcal{T}^n_h\right\}$
W	$= W^{1,\infty}(Q_T)$, S. 51
$\{\mathcal{B}_k\}_{k \ge 0}$	hierarchische modale Basis, S. 40
$\{\mathcal{B}_j\}_{1 \le j \le n^k_{dof}}$	eingebette hierarchische nodale Basis vom Grad K, S. 41
$V^{K,k}_h(T)$	$= \operatorname{span}\{\mathcal{B}_K\}$, S. 42
$V'_h(T)$	$= \operatorname{span}\{\mathcal{B}_k \setminus \mathcal{B}_K\}$, S. 42
F_T	$= J_T \hat{x} + b_T$ affine Transformation von $\hat{T}$ auf T
T^+	$= T$ Zeit-Raum Element
T^-	Nachbarelement von T^+ bzgl. der betrachteten Kante
$n^{\pm}_T$	Einheitsnormalenvektor bzgl. $\partial T^{\pm}$
$v^{\pm}(x)$	$= \lim_{\mu \to +0} v(x - \mu n^{\pm})$
$v^n_{\pm}(x)$	$= v(t_n \pm 0, x_1, \ldots, x_d)$

Literaturverzeichnis

[AF03] Adams, Robert A. und Fournier, John J. F. *Sobolev Spaces*. 2. Aufl. Department of Mathematics, the University of British Columbia, 2003. S. S. 8, 9, 12.

[AF75] Adams, Robert A. und Fournier, John J. F. *Sobolev Spaces*. 1. Aufl. Department of Mathematics, the University of British Columbia, 1975. S. S. 8.

[Alt02] Alt, Walter. *Nichtlineare Optimierung*. Braunschweig, Wiesbaden: Vieweg Verlag, 2002, S. 316. S. S. 62.

[Alt06] Alt, Hans Wilhelm. *Lineare Funktionalanalysis*. Berlin, Heidelberg: Springer-Verlag, 2006. S. S. 8–11, 13.

[Ang95] Angermann, Lutz. „Error estimates for the finite-element solution of an elliptic singularly perturbed problem“. In: *IMA J. Numer. Anal.* 15.2 (1995), S. 161–196. ISSN: 0272-4979. S. S. 97.

[Ape04] Apel, Thomas. „Interpolation in h-version Finite Element Spaces“. In: *Encyclopedia of computational mechanics. Vol. 1*. Hrsg. von Stein, Erwin, Borst, René de und Hughes, Thomas J. R. Bd. 1. Chichester: John Wiley & Sons Ltd., 2004, S. xii+798. S. S. 15.

[AS97] Ansorge, R. und Sonar, Th. „Informationsverlust, abstrakte Entropie und die mathematische Beschreibung des zweiten Hauptsatzes der Thermodynamik“. In: *Z. Angew. Math. Mech.* 77.11 (1997), S. 803–821. ISSN: 0044-2267. S. S. 2, 19.

[Bar98] Baran, Miroslaw. „New approach to Markov inequality in L^p norms“. In: *Approximation Theory*. Hrsg. von Govil, Narendra Kumar u. a. Monographs and textbooks in pure and applied mathematics. New York: Marcel Dekker, Inc., 1998, S. 75–85. S. S. 47.

[Bau62] Bauer, F. L. „On the field of values subordinate to a norm“. In: *Numerische Mathematik* 4 (1962), S. 103–113. ISSN: 0029-599X. S. S. 68.

[BB04] Becker, Roland und Braack, Malte. „A two-level stabilization scheme for the Navier-Stokes equations“. In: *Numerical mathematics and advanced applications*. Berlin: Springer, 2004, S. 123–130. S. S. 54.

[BD71] Bonsall, F. F. und Duncan, J. *Numerical ranges of operators on normed spaces and of elements of normed algebras*. Bd. 2. London Mathematical Society Lecture Note Series. London: Cambridge University Press, 1971, S. iv+142. S. S. 68.

[BHK] Bangerth, W., Hartmann, R. und Kanschat, G. „deal.II — a general-purpose object-oriented finite element library“. In: *ACM Trans. Math. Softw.* 33.4 (). S. S. 4.

[BL76] Bergh, Jöran und Löfström, Jörgen. *Interpolation spaces. An introduction*. Grundlehren der Mathematischen Wissenschaften, No. 223. Berlin: Springer-Verlag, 1976, S. x+207. S. S. 12, 16.

[BM97] Bernardi, Christine und Maday, Yvon. „Spectral Methods“. In: *Handbook of numerical analysis. Vol. V*. Hrsg. von Ciarlet, P. G. und Lions, J. L. Handbook of Numerical Analysis, V. Amsterdam: North-Holland, 1997, S. x+818. S. S. 39.

[BO04] Barth, Timothy und Ohlberger, Mario. „Finite Volume Methods: Foundation and Analysis“. In: *Encyclopedia of computational mechanics. Vol. 1*. Hrsg. von Stein, Erwin, Borst, René de und Hughes, Thomas J. R. Bd. 1. Chichester: John Wiley & Sons Ltd., 2004, S. xii+798. S. S. 52, 56.

[BRN79] Bardos, C., Roux, A. Y. le und Nédélec, J.-C. „First order quasilinear equations with boundary conditions“. In: *Communications in Partial Differential Equations* 4.9 (1979), S. 1017–1034. ISSN: 0360-5302. S. S. 20, 22.

[BS08] Brenner, Susanne C. und Scott, L. Ridgway. *The mathematical theory of finite element methods*. 3. Aufl. Bd. 15. Texts in Applied Mathematics. New York: Springer-Verlag, 2008, S. xviii+397. ISBN: 978-0-387-75933-3. S. S. 8.

[Cal+00] Calvetti, D. u. a. „Computation of Gauss-Kronrod quadrature rules“. In: *Math. Comp.* 69.231 (2000), S. 1035–1052. ISSN: 0025-5718. S. S. 42.

[Can+07] Canuto, C. u. a. *Spectral methods*. Scientific Computation. Evolution to complex geometries and applications to fluid dynamics. Berlin: Springer-Verlag, 2007, S. xxx+596. ISBN: 978-3-540-30727-3. S. S. 39.

[CB93] Chen, Q. und Babuška, I. *Polynomial interpolation of real functions I: Interpolation in an interval*. Technischer Bericht. College Park Campus, 1993. S. S. 46.

[Che01] Cheney, Ward. *Analysis for applied mathematics*. Bd. 208. Graduate Texts in Mathematics. New York: Springer-Verlag, 2001, S. viii+444. ISBN: 0-387-95279-9. S. S. 17.

[Cia91] Ciarlet, P.G. „Basic Error Estimates for Elliptic Problems“. In: *Handbook of numerical analysis. Vol. II*. Hrsg. von Ciarlet, P. G. und Lions, J.-L. Bd. II. Handbook of Numerical Analysis, II. Amsterdam: North-Holland, 1991, S. x+928. S. S. 15.

[CQ82] Canuto, C. und Quarteroni, A. „Approximation results for orthogonal polynomials in Sobolev spaces“. In: *Math. Comp.* 38.157 (1982), S. 67–86. ISSN: 0025-5718. S. S. 35, 40.

[Dix02] Dix, Daniel B. „Large-time behaviour of solutions of Burgers equation“. In: *Proceedings Section A: Mathematics - Royal Society of Edinburgh* 132 (2002). S. S. 13.

[DL93] DeVore, R. A. und Lorentz, G. G. *Constructive Approximation*. Bd. 303. Grundlehren der mathematischen Wissenschaften. Berlin: Springer-Verlag, 1993. S. S. 47.

[EG04] Ern, A. und Guermond, J. *Theory and practice of finite elements*. New York: Springer-Verlag, 2004. S. S. 11, 15, 27, 28, 31, 32, 36, 39, 46, 49–51, 54.

[EGK08] Eck, Christof, Garcke, Harald und Knabner, Peter. *Mathematische Modellierung (Springer-Lehrbuch) (German Edition)*. 1. Aufl. Springer, 2008. ISBN: 9783540749677. S. S. 2.

[Emm04] Emmerich, Etienne. *Gewöhnliche und Operator-Differentialgleichungen. Eine integrierte Einführung in Randwertprobleme und Evolutionsgleichungen für Studierende*. Wiesbaden: Vieweg Verlag, 2004, S. 300. S. S. 17.

[Fef00] Fefferman, Charles L. *Existence and smoothness of the Navier-Stokes equation*. 2000. URL: http://www.claymath.org/millennium/Navier-Stokes_Equations/. S. S. 1.

[Fej32] Fejér, Leopold. „Bestimmung derjenigen Abszissen eines Intervalles, für welche die Quadratsumme der Grundfunktionen der Lagrangeschen Interpolation im Intervalle ein Möglichst kleines Maximum Besitzt“. In: *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (2)* 1.3 (1932), S. 263–276. ISSN: 0391-173X. S. S. 40.

[Fri64] Friedman, Avner. *Partial differential equations of parabolic type*. Englewood Cliffs, NJ: Prentice Hall, 1964, S. xiv+346. S. S. 19.

[Fri96] Frisch, Uriel. *Turbulence: The Legacy of A. N. Kolmogorov*. Cambridge University Press, 1996. ISBN: 9780521457132. S. S. 1.

[Gas70] Gastinel, Noel. *Linear numerical analysis*. Translated from the original French text. Paris: Hermann, 1970, S. ix+341. S. S. 9.

[Geo03] Georgoulis, Emmanuil H. „Discontinuous Galerkin Methods on Shape-Regular and Anisotropic Meshes“. Dissertation. University of Oxford, 2003. S. S. 36.

[Giu84] Giusti, E. *Minimal Surfaces and Functions of Bounded Variation*. Bd. 80. Monographs in Mathematics. Boston: Birkhäuser, 1984. S. S. 8.

[GNP08] Gudi, Thirupathi, Nataraj, Neela und Pani, Amiya. „hp -Discontinuous Galerkin methods for strongly nonlinear elliptic boundary value problems“. In: *Numerische Mathematik* 109.2 (2008), S. 233–268. S. S. 46.

[Gri85] Grisvard, P. *Elliptic problems in nonsmooth domains*. Bd. 24. Monographs and studies in mathematics. Boston: Pitman Advanced Pub. Program, 1985, S. xiv+410. ISBN: 0-273-08647-2. S. S. 9, 11.

[Gue99] Guermond, Jean-Luc. „Stabilization of Galerkin approximations of transport equations by subgrid modeling“. In: *M2AN Math. Model. Numer. Anal.* 33.6 (1999), S. 1293–1316. ISSN: 0764-583X. S. S. 54.

[Gzy86] Gzyl, Henryk. „Multidimensional extension of Faa di Bruno's formula“. In: *J. Math. Anal. Appl.* 116.2 (1986), S. 450–455. ISSN: 0022-247X. S. S. 13.

[HB79] Hughes, T. J. R. und Brooks, A. „A multidimensional upwind scheme with no crosswind diffusion“. In: AMD 34 (1979), S. 19–35. S. S. 3.

[Hei07] Heitmann, Noel F. „Subgridscale stabilization of time dependent convection dominated diffusive transport“. In: *J. Math. Anal. Appl.* (2007). ISSN: 0022-247X. S. S. 4, 54, 99, 100.

[Heu92] Heuser, H. *Funktionalanalysis, Theorie und Anwendung*. Mathematische Leitfäden. Stuttgart, Leipzig: B. G. Teubner Verlag, 1992. S. S. 67.

[HFM86] Hughes, T. J. R., Franca, L. P. und Mallet, M. „A new finite element formulation for computational fluid dynamics. I. Symmetric forms of the compressible Euler and Navier-Stokes equations and the second law of thermodynamics“. In: *Comput. Methods Appl. Mech. Engrg.* 54.2 (1986), S. 223–234. ISSN: 0045-7825. S. S. 3.

[HFM87] Hughes, Thomas J. R., Franca, Leopoldo P. und Mallet, Michel. „A new finite element formulation for computational fluid dynamics. VI. Convergence analysis of the generalized SUPG formulation for linear time-dependent multidimensional advective-diffusive systems“. In: *Comput. Methods Appl. Mech. Engrg.* 63.1 (1987), S. 97–112. ISSN: 0045-7825. S. S. 3.

[HJS02] Houston, Paul, Jensen, Max und Süli, Endre. „hp-Discontinuous Galerkin Finite Element Methods with Least-Squares Stabilization“. In: *Journal of Scientific Computing* 17.1 (2002), S. 3–25. S. S. 95.

[HM86a] Hughes, Thomas J. R. und Mallet, Michel. „A new finite element formulation for computational fluid dynamics. III. The generalized streamline operator for multidimensional advective-diffusive systems“. In: *Comput. Methods Appl. Mech. Engrg.* 58.3 (1986), S. 305–328. ISSN: 0045-7825. S. S. 3.

[HM86b] Hughes, Thomas J. R. und Mallet, Michel. „A new finite element formulation for computational fluid dynamics. IV. A discontinuity-capturing operator for multidimensional advective-diffusive systems“. In: *Comput. Methods Appl. Mech. Engrg.* 58.3 (1986), S. 329–336. ISSN: 0045-7825. S. S. 3.

[HMM86] Hughes, Thomas J. R., Mallet, Michel und Mizukami, Akira. „A new finite element formulation for computational fluid dynamics. II. Beyond SUPG“. In: *Comput. Methods Appl. Mech. Engrg.* 54.3 (1986), S. 341–355. ISSN: 0045-7825. S. S. 3.

[HST37] Hille, Einar, Szegö, G. und Tamarkin, J. D. „On some generalizations of a theorem of A. Markoff“. In: *Duke Math. J.* 3.4 (1937), S. 729–739. ISSN: 0012-7094. S. S. 46, 47.

[HT84] Hughes, T. J. R. und Tezduyar, T. E. „Finite element methods for first-order hyperbolic systems with particular emphasis on the compressible Euler equations“. In: *Comput. Methods Appl. Mech. Engrg.* 45.1-3 (1984), S. 217–284. ISSN: 0045-7825. S. S. 3.

[JJS95] Jaffré, J., Johnson, C. und Szepessy, A. „Convergence of the discontinuous Galerkin finite element method for hyperbolic conservation laws“. In: *Math. Models Methods Appl. Sci.* 5.3 (1995), S. 367–386. ISSN: 0218-2025. S. S. iii, 3, 4, 53, 54, 81, 92, 99.

[JK07] John, Volker und Knobloch, Petr. „On spurious oscillations at layers diminishing (SOLD) methods for convection-diffusion equations. I. A review“. In: *Comput. Methods Appl. Mech. Engrg.* 196.17-20 (2007), S. 2197–2215. ISSN: 0045-7825. S. S. 53.

[JK08a] John, Volker und Kaya, Songul. „Finite element error analysis of a variational multiscale method for the Navier-Stokes equations“. In: *Advances in Computational Mathematics* 28.1 (2008), S. 43–61. S. S. 54.

[JK08b] John, Volker und Knobloch, Petr. „On spurious oscillations at layers diminishing (SOLD) methods for convection-diffusion equations. II. Analysis for Pb_x1 and Qb_x1 finite elements“. In: *Comput. Methods Appl. Mech. Engrg.* 197.21-24 (2008), S. 1997–2014. ISSN: 0045-7825. S. S. 53.

[JKL06] John, Volker, Kaya, Songul und Layton, William. „A two-level variational multiscale method for convection-dominated convection-diffusion equations“. In: *Comput. Methods Appl. Mech. Engrg.* 195.33-36 (2006), S. 4594–4603. ISSN: 0045-7825. S. S. 54.

[Joh87] Johnson, Claes. *Numerical solution of partial differential equations by the finite element method*. Cambridge: Cambridge University Press, 1987, S. 278. ISBN: 0-521-34514-6; 0-521-34758-0. S. S. 3, 51.

[JP86] Johnson, C. und Pitkäranta, J. „An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation“. In: *Math. Comp.* 46.173 (1986), S. 1–26. ISSN: 0025-5718. S. S. 3, 51, 53.

[JS86a] Johnson, C. und Szepessy, A. „On the convergence of streamline diffusion finite element methods for hyperbolic conservation laws“. In: *Numerical methods for compressible flows - Finite difference, element and volume techniques; Proceedings of the Winter Annual Meeting, Anaheim, CA, Dec. 7-12, 1986 (A87-38490 16-02). New York, American Society of Mechanical Engineers, 1986, p. 75-91.* Bd. 7. 1986, S. 75–91. S. S. 3.

[JS86b] Johnson, Claes und Saranen, Jukka. „Streamline diffusion methods for the incompressible Euler and Navier-Stokes equations“. In: *Math. Comp.* 47.175 (1986), S. 1–18. ISSN: 0025-5718. S. S. 56.

[JS87] Johnson, Claes und Szepessy, Anders. „On the convergence of a finite element method for a nonlinear hyperbolic conservation law“. In: *Math. Comp.* 49.180 (1987), S. 427–444. ISSN: 0025-5718. S. S. 3.

[JSH90] Johnson, Claes, Szepessy, Anders und Hansbo, Peter. „On the convergence of shock-capturing streamline diffusion finite element methods for hyperbolic conservation laws“. In: *Math. Comp.* 54.189 (1990), S. 107–129. ISSN: 0025-5718. S. S. 3, 99.

[KL09] Knobloch, Petr und Lube, Gert. „Local projection stabilization for advection-diffusion-reaction problems: One-level vs. two-level approach“. In: *Applied Numerical Mathematics* 59.12 (2009), S. 2891 –2907. ISSN: 0168-9274. S. S. 54.

[KR05] Kaya, Songul und Rivière, Béatrice. „A discontinuous subgrid eddy viscosity method for the time-dependent Navier-Stokes equations“. In: *SIAM J. Numer. Anal.* 43.4 (2005), 1572–1595 (electronic). ISSN: 0036-1429. S. S. 54.

[KS05] Karniadakis, George Em und Sherwin, Spencer J. *Spectral/hp element methods for computational fluid dynamics*. 2. Aufl. Numerical Mathematics and Scientific Computation. New York: Oxford University Press, 2005, S. xviii+657. ISBN: 978-0-19-852869-2; 0-19-852869-8. S. S. 34.

[LS04] Li, Chi-Kwong und Sourour, Ahmed Ramzi. „Linear operators on matrix algebras that preserve the numerical range, numerical radius or the states“. In: *Canad. J. Math.* 56.1 (2004), S. 134–167. ISSN: 0008-414X. S. S. 68.

[Mal+96] Malek, J. u. a. *Weak and Measure-valued Solutions to Evolutionary PDEs*. Bd. 13. Applied Mathematics and Mathematical Computation. Chapman & Hall, 1996, S. xi + 317. S. S. 23, 24, 26.

[MMR94] Milovanovic, G. V., Mitrinovic, D. S. und Rassias, Th. M. *Topics in Polynomials: Extremal Problems, Inequalities, Zeros*. Singapore: World Scientific Publ. Co, 1994, S. XIV+822. S. S. 47.

[MST07] Matthies, Gunar, Skrzypacz, Piotr und Tobiska, Lutz. „A unified convergence analysis for local projection stabilisations applied to the Oseen problem“. In: *M2AN Math. Model. Numer. Anal.* 41.4 (2007), S. 713–742. ISSN: 0764-583X. S. S. 54.

[MT98] Marion, Martine und Temam, Roger. „Navier-Stokes Equations: Theory and Approximation“. In: *Handbook of numerical analysis. Vol. VI*. North-Holland, 1998. S. S. 2.

[Nik51] Nikolskii, S. M. „Inequalities for entire functions of finite degree and their application in the theory of differentiable functions of several variables“. In: *Trudy Mat. Inst. Steklov., v. 38*. Trudy Mat. Inst. Steklov., v. 38. Moscow: Izdat. Akad. Nauk SSSR, 1951, S. 244–278. S. S. 46.

[NR50] Neumann, J. von und Richtmyer, R. D. „A Method for the Numerical Calculation of Hydrodynamic Shocks“. In: *Journal of Applied Physics* 21.3 (1950), S. 232–237. S. S. 3.

[NS64] Nirschl, Nicholas und Schneider, Hans. „The Bauer fields of values of a matrix“. In: *Numer. Math.* 6 (1964), S. 355–365. ISSN: 0029-599X. S. S. 68.

[QV94] Quarteroni, Alfio und Valli, Alberto. *Numerical approximation of partial differential equations*. Bd. 23. Springer Series in Computational Mathematics. Berlin: Springer-Verlag, 1994, S. xvi+543. ISBN: 3-540-57111-6. S. S. 17, 34.

[RH73] Reed, W.H. und Hill, T.R. *Triangular mesh methods for the neutron transport equation*. Technischer Bericht LA-UR-73-479. Los Alamos, New Mexico: Los Alamos National Laboratory, 1973. S. S. 51.

[Sma63] Smagorinsky, J. „GENERAL CIRCULATION EXPERIMENTS WITH THE PRIMITIVE EQUATIONS“. In: *Monthly Weather Review* 91.3 (März 1963), S. 99–164. S. S. 3.

[Ste70] Stein, Elias M. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton, N.J.: Princeton University Press, 1970, S. xiv+290. S. S. 12.

[SVD06] Sudirham, J. J., Vegt, J. J. W. van der und Damme, R. M. J. van. „Space-time discontinuous Galerkin method for advection-diffusion problems on time-dependent domains“. In: *Appl. Numer. Math.* 56.12 (2006), S. 1491–1518. ISSN: 0168-9274. S. S. 36.

[Sze89a] Szepessy, Anders. „Convergence of a shock-capturing streamline diffusion finite element method for a scalar conservation law in two space dimensions“. In: *Math. Comp.* 53.188 (1989), S. 527–545. ISSN: 0025-5718. S. S. 3, 4, 53, 67, 81, 99.

[Sze89b] Szepessy, Anders. „Measure-valued solutions of scalar conservation laws with boundary conditions“. In: *Arch. Rational Mech. Anal.* 107.2 (1989), S. 181–193. ISSN: 0003-9527. S. S. 100.

[Sze91] Szepessy, A. „Convergence of a streamline diffusion finite element method for scalar conservation laws with boundary conditions“. In: *RAIRO Modél. Math. Anal. Numér.* 25.6 (1991), S. 749–782. ISSN: 0764-583X. S. S. 3, 53, 67, 79, 99.

[Tim63] Timan, A. F. *Theory of approximation of functions of a real variable*. Translated from the Russian by J. Berry. English translation edited and editorial preface by J. Cossar. International Series of Monographs in Pure and Applied Mathematics, Vol. 34. A Pergamon Press Book. The Macmillan Co., New York, 1963, S. xii+631. S. S. 46.

[Toe18] Toeplitz, Otto. „Das algebraische Analogon zu einem Satze von Fejér“. In: *Math. Z.* 2.1-2 (1918), S. 187–197. ISSN: 0025-5874. S. S. 68.

[Tri78] Triebel, Hans. *Interpolation theory, function spaces, differential operators*. Bd. 18. North-Holland Mathematical Library. Amsterdam: North-Holland Publishing Co., 1978, S. 528. ISBN: 0-7204-0710-9. S. S. 16.

[Vov02] Vovelle, Julien. „Convergence of finite volume monotone schemes for scalar conservation laws on bounded domains“. In: *Numerische Mathematik* 90.3 (2002), S. 563–596. S. S. 26.

[War99] Warnecke, Gerald. *Analytische Methoden in der Theorie der Erhaltungsgleichungen*. Bd. 138. Teubner-Texte zur Mathematik. Stuttgart, Leipzig: B. G. Teubner, 1999, S. 344. S. S. 8, 15, 19.

[Wlo82] Wloka, Joseph. *Partielle Differentialgleichungen*. Mathematische Leitfaeden. Stuttgart: B. G. Teubner, 1982, S. 500. S. S. 8, 11.

[ZF84] Zurmühl, Rudolf und Falk, Sigurd. *Matrizen und ihre Anwendungen für angewandte Mathematiker, Physiker und Ingenieure. Teil 1*. 5. Aufl. Grundlagen. Berlin: Springer-Verlag, 1984, S. xiv+342. ISBN: 3-540-12848-4. S. S. 70.